

# La regla de oro del utilitarismo

### Un refrán conocido

Si usted discute sobre temas morales (especialmente con un lego), lo más probable es que en los primeros cinco minutos de la discusión su contrincante la invoque, en alguna de sus versiones. Pero si usted le pide a un estudiante de filosofía moral que la rebata, lo más probable es que demore menos de cinco minutos en hacerlo. ¿De qué se trata, en este caso?: de la Regla de Oro, por supuesto (desde aquí, RO). Pero ¿cómo es posible que una regla tan vulnerable tenga –a la vez– tanta influencia? Y no solo entre los legos, por cierto, ya que los filósofos también a veces recurren a ella.

Los partidarios de la RO asumen implícitamente que existen similitudes importantes entre nosotros y los demás agentes morales<sup>1</sup> y que estas similitudes tornan plausible la RO. Frente a esta postura, los adversarios de la RO sugieren que *no* hagamos a los demás lo que querríamos que nos hicieran a nosotros, porque sus gustos pueden ser diferentes (como lo hizo notar Bernard Shaw, en el Prefacio de *Hombre y Superhombre*). Pero este rasgo por sí solo no alcanza para dar cuenta de la profundidad de las diferencias entre los defensores y los detractores de la RO.

Filosóficamente hablando, en un extremo tenemos a Kant, que descarta sumariamente a la RO, y en el otro a Mill, que la recomienda con entusiasmo. ¿Es la RO, entonces, un instrumento exclusivamente utilitarista?

Estos son algunos de los problemas que rodean al tema, y voy a proceder a estudiarlos de la forma que ahora enunciaré. En la sección 1, voy a considerar las dos maneras más populares de enunciar la RO y a explicar sus diferencias. En la sección 2, me referiré brevemente al aspecto religioso de la RO, pero de una

---

<sup>1</sup> JEFFREY WATTLES, *The Golden Rule*, Oxford University Press, 1996, p. 115.

forma negativa: quiero explicitar que este es el aspecto del cual *no* voy a ocuparme. Examino también allí la posición de Freud acerca de la RO religiosa. En la sección 3, muestro que la RO es algo parecido a un concepto y no una concepción, y en la sección 4, identifico a la consistencia como la prescripción central de la RO, tratando de deflectar algunas de las críticas más usuales a la misma. La sección 5 se concentra en el ámbito de aplicación de la RO, y la 6, en la relación entre la RO y la motivación. En la sección 7, estudio la actitud de Kant frente a la RO, y en la sección 8, la actitud de Mill y la de Sidgwick frente a la misma. La 9 explica el papel de la RO dentro del utilitarismo, y la sección 10 concentra mis conclusiones.

## Formulación

La RO puede enunciarse tanto dirigida a una acción, como dirigida a una omisión. Entendida en términos de acción, la RO dice: “Trata a los demás como desearías que te trataran a ti”. Entendida en términos de omisión, la RO dice: “No trates a los demás como no desearías que te trataran a ti”. Voy a mostrar que la RO es mucho más exigente si se la entiende como requiriendo una acción que si se la entiende como demandando una omisión.

Autores como Marcus Singer mencionan un argumento que sostiene que las dos formulaciones son equivalentes, puesto que cualquier acción puede ser descripta en términos positivos o negativos: decir “A quiere que pase *x*” es equivalente a decir “A no quiere que no pase *x*”. Cualquier deseo formulado negativamente puede ser reformulado positivamente, de modo que no existe ninguna diferencia lógica o moral entre las dos formulaciones, sino solo una diferencia psicológica o retórica.<sup>2</sup>

No cuestiono este punto, pero mi afirmación acerca del grado de exigencia de la RO apunta a otra cosa. Decir “Haz a los demás...” requiere del agente una acción, mientras que decir “No hagas a los demás...” requiere del agente una omisión. Requerir una acción implica más exigencia que requerir una omisión, puesto que en el primer caso el agente tiene menos libertad que en el segundo. Veamos por qué.

Supongamos estas dos proposiciones: a) “A quiere que B use corbata azul”, y b) “A no quiere que B use corbata azul”. Lo que hace Singer es mostrar que a) equivale a esta proposición: “A no quiere que B no use corbata azul”, pero esta equivalencia no es relevante para medir lo que a) exige del agente respecto de lo que b) exige. En el caso

---

<sup>2</sup> MARCUS G. SINGER, “Golden Rule”, en Paul Edwards (ed.), *The Encyclopedia of Philosophy*, New York, MacMillan Publishing Co., 1967, vol.3, pp. 365-366.

a), el agente B no tiene más que una alternativa (suponiendo que él debe obedecer a A): vestirse con una corbata azul. En el caso b), en cambio, B tiene varias alternativas disponibles, tales como las de usar una corbata verde o amarilla, o no usar corbata.

Si estoy obligado a hacer  $x$ , tengo menos libertad que si estoy obligado a no hacer  $x$ , porque en este segundo caso todavía puedo hacer muchas otras cosas. El agente B, entonces, es más libre en el segundo caso que en el primero. En definitiva, los deberes positivos son más exigentes que los deberes negativos, y no creo que nadie quiera discutir esto.

Luego veremos que la RO en su versión utilitarista se enuncia como requiriendo una acción, lo cual es compatible con las exigencias morales del utilitarismo.

## El aspecto religioso

El judaísmo suele prescribir la RO como una omisión y el cristianismo suele hacerlo como una acción.<sup>3</sup> Ese aspecto de la distinción no va a ocuparme aquí, sin embargo, puesto que voy a examinar la RO desde la óptica moral y no religiosa.

Las diferencias entre ambas posturas son harto evidentes, y las frecuentes confusiones entre ambas se deben a la circunstancia de que las religiones predicán también su propia moral, y los fieles suelen identificar ambas cosas. Pero la religión, por una parte, es una cuestión de fe, y sus enunciados no requieren de argumento alguno, puesto que son considerados como dogmas. El concepto de *pecado*, por ejemplo, se asocia con el de alguna acción –u omisión– que disgusta –o entristece– a Dios. El creyente no tiene pruebas de ese disgusto o tristeza eventual, pero tiene fe en lo que le dicen sus predicadores. El creyente no necesita entonces conocer ninguna razón por la que Dios deba entristecerse o disgustarse, solo necesita saber cuáles son las cosas que le producen a Dios esa reacción. Supongamos, por ejemplo, que los predicadores enseñan que a Dios le disgustan –o entristecen– las relaciones sexuales fuera del matrimonio, incluso entre adultos normales que consienten y que son legalmente libres para hacerlo. El creyente, entonces, debe abstenerse de practicar tales relaciones, porque ellas son pecaminosas.

Pero ¿son ellas inmorales? Esta es una pregunta muy distinta, porque para responderla debemos proporcionar argumentos y razones. Si preguntamos por qué esas relaciones pueden ser inmorales no nos ayuda en nada decir que ellas disgustan –o entristecen– a Dios. Dado que en muchos casos no existen argumentos en

<sup>3</sup> HARRY J. GENSLER, *Formal Ethics*, London, Routledge, 2007, p. 105.

contra de ese tipo de relaciones, en tales casos ellas pueden ser pecaminosas, pero ciertamente no son inmorales. (El capítulo 10 profundiza este tema).

A mí me interesa únicamente la dimensión moral de la RO y no su dimensión religiosa. Solo me ocuparé de esta última para mostrar cómo la RO suele aquí requerir una formulación distinta, que ha sido justificadamente criticada.

La formulación usual de la RO apela explícitamente a la reciprocidad e implícitamente a la simpatía. Algunos autores<sup>4</sup> parten de estos supuestos, pero también sugieren que la simpatía puede tener corto alcance y no ser una guía suficiente, de donde la suplementan con la razón, puesto que la conducta debe ser a la vez bienintencionada e inteligente, y concluyen luego llevando la RO al plano religioso y sugiriendo que ella “es el principio de la práctica de la familia de Dios”.<sup>5</sup>

Ninguna religión en particular recurre a la RO, sino que todas lo hacen. Por eso Wilson sostiene que ella captura la esencia de la religión en tanto hermandad universal. Pero si la RO aparece en todas las religiones, por supuesto no puede esperarse una sola versión de ella, sino una gran cantidad de versiones, a la par que debe aceptarse que este rasgo de la RO indica que ella no puede explicar la naturaleza específica de ninguna religión en particular.<sup>6</sup>

Como he dicho, las versiones seculares de la RO apelan a la reciprocidad; las versiones religiosas enfatizan, en cambio, el amor: “Ama al prójimo como a ti mismo”. Esta es la versión de la RO que Freud critica en *El malestar en la cultura* por considerar que la regla prescinde de la psicología, que ella no es psicológicamente realista; por eso, Freud pregunta cómo podríamos llegar a cumplir con este precepto, de qué manera podríamos satisfacerlo. Porque el amor no depende de una decisión deliberada: alguien puede ordenarme que me comporte *como* si sintiera amor por alguien, que trate a los demás *como* si los amara, pero no puede ordenarme que los ame.

La crítica freudiana, no obstante, es todavía más profunda, y cuestiona también el contenido mismo de la RO religiosa:

*Si amo a alguien, es preciso que este lo merezca por cualquier título (...).*

*Merecería mi amor si se me asemejara en aspectos importantes (...); lo*

*merecería si fuera más perfecto de lo que yo soy (...). En cambio, si me*

<sup>4</sup> JEFFREY WATTLES, “Philosophical Reflections on the Golden Rule”, en *Simposio The Golden Rule in Religions of the World*, Bard’s Institute of Advanced Theology. Consultar en [www.bard.edu/iat/](http://www.bard.edu/iat/)

<sup>5</sup> WATTLES, “Philosophical Reflections...”, cit., pp. 5, 7 y 13.

<sup>6</sup> DAVID SLOAN WILSON, “The Golden Rules of Religions”, en el Simposio cit., pp. 3,4,10 y 13.

*fuera extraño y si no me atrajese ninguno de sus propios valores (...) entonces me sería muy difícil amarlo. Hasta sería injusto si lo amara, pues los míos aprecian mi amor como una demostración de preferencia, y les haría injusticia si los equiparara con un extraño.*

De modo que Freud formula dos objeciones a esta versión de la RO religiosa: a) ella no se puede cumplir, y b) aunque pudiera cumplirse, muchas veces sería injusta. Pero existe otra versión compatible con la RO religiosa que Freud no cuestiona, y es la que dice “Amarás al prójimo como el prójimo te ama a ti”, respecto de la cual él reconoce que no tiene nada que objetar.<sup>7</sup>

Por otra parte, no todas las versiones religiosas de la RO invocan solo el amor. Samuel Clarke, por caso, luego de exigir la honra de Dios, le asigna a la RO esta formulación: nuestro deber consiste en “tratar con cualquier hombre como, en circunstancias similares, podríamos esperar razonablemente que él tratara con nosotros”. Clarke no está invocando aquí el amor, ni tampoco la reciprocidad, ni la simpatía: él recurre en cambio a la equidad como sustento de la RO, a un deber de juego limpio.<sup>8</sup> Y piensa que hay incluso algo de verdad analítica en su planteo, porque

*cualquier cosa que yo considere razonable o irrazonable que otro me haga a mí; eso, por el mismo juicio, yo declaro razonable o irrazonable, que yo le hiciera a él en un caso similar. Y negar esto en palabra o en acción, sería como si un hombre pretendiera que, aunque dos y tres es igual a cinco, sin embargo cinco no es igual a dos y tres.<sup>9</sup>*

Aunque reconozco alguna intuitividad en la presentación de Clarke, pienso que su paralelismo con la aritmética es claramente exagerado, e insistiré luego en que la RO no es autoevidente, ni –por supuesto– analíticamente verdadera.

Pero, como dije, la religión queda fuera de mis propósitos, por lo que vuelvo a la moral y a formulaciones de la RO que no puedan ser víctimas de la crítica freudiana.

<sup>7</sup> SIGMUND FREUD, “El malestar en la cultura”, en *Obras Completas*, Buenos Aires, Editorial Losada, 1997, tomo 22, pp. 3044-3045.

<sup>8</sup> Invoca enseguida el amor, no obstante, cuando exige que nos esforcemos para promover el bienestar y la felicidad de todos los hombres.

<sup>9</sup> SAMUEL CLARKE, “A Discourse of Natural Religion”, en D. D. Raphael (ed.), *British Moralists 1650/1800*, Oxford, Clarendon Press, 1969, vol. I pp. 207-208.

## Conceptos y concepciones

Dentro del ámbito de la moral, un reproche que se le dirige frecuentemente a la RO es que ella es demasiado abstracta y vaga, y el reproche es acertado aunque inevitable, puesto que la RO es algo semejante a un concepto y no una concepción. Cuando Kant le reprocha vaguedad a la RO, está olvidando que se enfrenta a un (cierto tipo de) concepto.

El concepto se distingue de sus posibles concepciones, precisamente, por ser abstracto y vago, compatible –así– con diversas concepciones. “Dar a cada uno lo suyo” es un concepto de justicia; determinar qué es lo suyo de cada uno da lugar a distintas concepciones de justicia. En base al concepto no podemos profetizar cuál va ser la distribución del ingreso y la riqueza en una sociedad; en base a una concepción, en cambio, sí podemos hacerlo. Varios modos de distribuir el ingreso y la riqueza son compatibles con el concepto de dar a cada uno lo suyo.

Algo parecido ocurre con la RO. Tratar a los demás como uno querría ser tratado es un (cierto tipo de) concepto; en base a él no podemos predecir específicamente la forma en que las personas van a ser tratadas, puesto que para hacerlo debemos pasar del concepto a la concepción. La concepción precisa la forma específica que deben adoptar los deseos del agente, y la predicción se torna –así– posible.

En su versión religiosa, la RO arranca, como corresponde, adoptando la forma de un concepto: “Ama al prójimo como a ti mismo”. Esta formulación torna imposible toda predicción, puesto que no sabemos cómo se ama el agente a sí mismo. Pero Jesús<sup>10</sup> lo transforma luego en una concepción: “Como yo los he amado, ustedes deben amarse los unos a los otros”. Ahora la RO tiene un contenido adicional y tenemos una idea más precisa respecto del tipo de amor que es requerido: hay que amar a los demás de la forma en que Jesús los amó. En este caso, las predicciones acerca de la conducta futura del agente se vuelven posibles. De todas formas, digo esto solo con propósitos ilustrativos, ya que no estoy sosteniendo que la RO sea *exactamente* un concepto; los ejemplos anteriores solo tienden a explicar su funcionamiento. Y su funcionamiento indica que no se pueden inferir de ella deberes concretos, sino que la RO requiere operar en combinación con otros principios, como luego veremos.

---

<sup>10</sup> En San Juan, 13:34 y 15:12.

## Consistencia y algo más

Pero si la RO es un tipo de concepto y –como tal– no proporciona respuestas morales específicas, ¿cuál puede ser su utilidad, más allá de la utilidad derivada de inspirar distintas concepciones? La RO demanda consistencia en nuestra conducta moral: ella puede ser entendida –entonces– como una condición necesaria para que una acción sea moral, pero por cierto que no es una condición suficiente. La RO puntualiza ciertas eventuales inconsistencias en nuestro comportamiento, inconsistencias que suelen provenir de descartar los intereses de los demás; trabaja entonces mejor en equipo, conjuntamente con otras demandas morales, y su papel central consiste en prevenirnos contra el autocentrismo.<sup>11</sup>

Se puede –así– satisfacer las exigencias de la RO y –a la vez– actuar de una manera que consideraríamos moralmente incorrecta, ya que podemos juzgar que los deseos del agente son moralmente defectuosos. Tomemos como ejemplo el caso que Hare analiza en *Freedom and Reason*, donde se pregunta seriamente si corresponde llamar moral a la posición del fanático nazi que acepta ser exterminado en el supuesto de que se pruebe que él mismo es judío.<sup>12</sup> El fanático nazi respeta las exigencias de la RO y, a la vez, exhibe una conducta inmoral.

Todo esto puede conducirnos a la tentación de complementar desde un comienzo la formulación de la RO. Algunos autores<sup>13</sup> sugieren agregarle a la consistencia el respeto por las personas, tratar a los demás con respeto, tal como quisiéramos nosotros ser tratados con respeto. O incorporar la idea de que la RO expresa un compromiso de tratar a los demás con consideración y equidad, ya que la RO debe presuponer un cierto nivel de madurez moral en el agente.<sup>14</sup>

En la misma línea que estoy exponiendo, San Agustín la formula de este modo: “Todas aquellas buenas cosas, en consecuencia, que quisieras que los hombres te hicieran a ti, hacéselas también a ellos”. La idea de San Agustín es –en principio– sensata. Él recuerda que la versión griega de la RO no contiene la referencia a lo bueno, y que tal referencia fue agregada en la versión latina para aclarar su significado. Porque

*si alguien desea que le hagan alguna cosa mala y para este propósito invocara este texto; por ejemplo, si una persona deseara ser desafiada*

<sup>11</sup> HARRY J. GENSLER, “Gold or Fool’s Gold?. Ridding the Golden Rule of Absurd Implications”, en el Simposio citado, pp. 8 y 14.

<sup>12</sup> R. M. HARE, *Freedom and Reason*, Oxford University Press, 1978, pp. 157-185.

<sup>13</sup> WATTLES, *The Golden Rule*, cit., p. 139.

<sup>14</sup> WATTLES, “Philosophical Reflections...”, cit., pp. 2 y 11.

*a beber inmoderadamente (...) y primero practicara esto sobre la persona por quien él desea ser desafiada: sería entonces ridículo que ese individuo supusiera que ha estado a la altura de la prescripción.*<sup>15</sup>

Aunque parece rechazar el criterio agustiniano, Gewirth emplea un método similar, pues sostiene que los deseos mencionados en la RO deben ser racionales, de donde la regla diría “Haz a los demás como querrías racionalmente que ellos te hicieran a tí”. La exigencia de la racionalidad puede interpretarse de dos maneras distintas, esto es, de un modo moralmente normativo y de un modo moralmente neutral. Gewirth rechaza la primera alternativa —que es la alternativa estrictamente agustiniana— porque piensa que en este caso uno ya sabe lo que es moralmente correcto, mientras que la RO pretende establecer ella misma el criterio de corrección moral.

Según Gewirth, la segunda alternativa recurre solo a los criterios de la lógica deductiva e inductiva, pero —curiosamente— de estos criterios él deduce un contenido necesario respecto de la RO, contenido que implica el derecho a la libertad y al bienestar: el agente debe abstenerse de interferir con la libertad y con el bienestar de todas las demás personas; partiendo de la lógica, Gewirth concluye sin embargo obteniendo un principio moral normativo.<sup>16</sup> Desgraciadamente, entonces, *racional* cumple en el intento de Gewirth el mismo papel que *bueno* cumple en el intento de San Agustín.

Si bien todos estos complementos evitan, desde luego, muchas de las críticas a las que la RO ha sido sometida, lo hacen a costa de incorporar a ella un criterio independiente de bondad o de corrección moral; la RO deja así de ser un concepto y se transforma en una de sus posibles concepciones.<sup>17</sup> Varias de las críticas a la RO, sin embargo, pueden ser rebatidas sin recurrir a estas adiciones. Veamos dos ejemplos muy comunes. a) Muchos de los casos en los cuales la RO parece proporcionar una respuesta absurda son casos en los que se pasan por alto las diferentes circunstancias en las que se encuentran los agentes involucrados: A es sordo y B debería hablarle en voz muy alta, pero B no quiere que le hablen en voz muy alta a él. Prescindir de este detalle —lo que resulta absurdo— es lo que lleva a la objeción burlona de Bernard Shaw a la RO que he mencionado antes. b) Sin invocar valores tales como la equidad

<sup>15</sup> St. Augustine, *The Lord's Sermon of the Mountain*, Westminster, MD, The Newman Press, 1948, p. 161.

<sup>16</sup> Cfr. ALAN GEWIRTH, “The Golden Rule Rationalized”, en Peter French, Theodor Uehling & Howard Wettstein (eds.), *Midwest Studies in Philosophy*, vol.III, Minneapolis, Minnesota University Press, 1980, pp. 137-140.

<sup>17</sup> Al par de no servir ya para identificar conductas morales: las cosas buenas ya las conozco sin recurrir a la RO. Es lo mismo que ocurre con el principio milliano del daño, en el cual no puede identificarse una conducta como dañina sin referirse al mismo tiempo a intereses protegidos por el derecho.



o la bondad, la RO puede robustecerse frente a las críticas incorporando la idea del consentimiento, y enunciándola de este modo: “Trata a los demás solo como tú consentirías ser tratado en la misma situación”.<sup>18</sup> No estoy sugiriendo, desde luego, que estas versiones resistan todas las críticas, pero sí que la RO no es tan débil como sus críticos sugieren. (Aunque –en definitiva– propondré en su momento que ella se emplee complementando un principio que especifique la idea de lo bueno).

Lo que me parece importante que se tenga presente es que la RO debe entenderse como referida a nuestra reacción presente frente a un caso hipotético, pero no como referida a la forma en que reaccionaríamos si estuviéramos ahora involucrados en el caso. Supongamos que un padre quiere aplicar la RO para decidir si va o no a pegarle a su hijo como consecuencia de su mal comportamiento. El padre debe preguntarse a sí mismo “¿Estoy dispuesto a ser golpeado si estuviera en el lugar de mi hijo?”, pero nunca debería preguntarse “¿Si yo estuviera ahora en el lugar de mi hijo, estaría dispuesto a ser golpeado?”. Si se formulara la segunda pregunta, concluiría erróneamente que nunca debería pegarle a su hijo, con independencia de su comportamiento.<sup>19</sup> Luego veremos que este es el error que cometió Kant al aplicar la RO al caso del juez que debe condenar a un delincuente.

## Ámbito de aplicación: el individuo y la sociedad

¿A quién está dirigida la RO? A primera vista, podríamos contestar confiadamente que se dirige al individuo, y que le prescribe un tipo de comportamiento moral. Pero es un tipo de comportamiento moral tal que necesariamente requiere la presencia de otros individuos, ya que la RO nos dice cómo debemos tratar a *los demás*.<sup>20</sup> La RO, entonces, asume la existencia de otros individuos, además del agente moral al que la regla parece dirigirse.

Consideremos el caso de Hobbes. Él enuncia la RO cuando se refiere al principio de la máxima libertad:

*Que un hombre (...) se contente con tanta libertad contra otro hombre como consentiría a otro hombre contra él mismo.*

<sup>18</sup> GENSLER, “Gold or...”, cit., pp. 2 y 4.

<sup>19</sup> GENSLER, “Gold or...”, cit., p. 6.

<sup>20</sup> La RO no especifica claramente quiénes son *los demás*: si ellos son todos los agentes morales, podríamos terminar directamente en el utilitarismo, sin intermediarios.

Y, desde luego, parece dirigir la RO al individuo. Pero el lector de Hobbes no puede dejar de advertir que —al mismo tiempo— se está allí enunciando un principio necesario para vivir en la sociedad civil. Un hombre debe renunciar a su derecho a toda cosa en favor de la paz, siempre que los demás también lo hagan. Esta, dice Hobbes,

*es aquella ley del Evangelio: todo aquello que requerís que otros os hagan, hacédselo a ellos, y aquella ley de todo hombre, quod tibi fieri non vis, alteri non facis.*<sup>21</sup>

Yo voy a entenderla, pues, de ambas formas: la RO se dirige al individuo, pero es un requerimiento para vivir en sociedad. En rigor, si la describimos en términos de reciprocidad, la RO parece no tener límites, no dice que debemos tratar de una cierta manera a los miembros de nuestra familia, o de nuestra sociedad, sino que dice que debemos tratar de una cierta manera a *todos* aquellos con los que podamos interactuar.

Pero esta característica de la RO nos revela un aspecto interesante de ella: puesto que necesariamente requiere la presencia de otros agentes, la RO no se refiere a deberes respecto de nosotros mismos. Si la RO se considera un resumen de la moral, entonces la moral se entiende como un conjunto de deberes respecto de terceros, y no de comportamientos autorreferentes.<sup>22</sup>

Este aspecto de la RO merecería —sin duda— la aprobación de Mill, puesto que él pensaba que la moral se reducía a nuestras relaciones con terceros, y no incluía deberes morales del agente para consigo mismo. Tal vez el motivo de Mill para establecer este límite residía en la circunstancia de que él deseaba acotar la intervención del derecho —especialmente del derecho penal— en la vida privada de los ciudadanos, y sostenía que el único motivo para aplicar un castigo residía en haber causado un daño a un tercero sin su consentimiento. Él identificaba la responsabilidad con el castigo: cuando se dice que sentimos ser moralmente responsables por nuestras acciones, la principal idea en la mente del que habla es la de ser castigado. Complementemos ahora esta afirmación con el principio milliano del daño, según el cual solo puede castigarse el daño no consentido a un

<sup>21</sup> THOMAS HOBBS, *Leviatán*, Madrid, Editora Nacional, 1979, parte primera, capítulo XIV.

<sup>22</sup> Hoche piensa que la RO puede generar deberes para con uno mismo. Sostiene que ella "menciona solo una actitud negativa respecto de la acción de otros, sin que tengamos necesariamente que presuponer que estas acciones afectan a alguien más". HANS-ULRICH HOCHÉ, "The Golden Rule: New Aspects of an Old Moral Principle", en Darrel E. Christensen (ed.), *Contemporary German Philosophy*, The Pennsylvania State University Press, 1982, vol. I, p. 84. Esta interpretación no me convence.

tercero, y se sigue que la conducta autorreferente no puede ser exigida, de donde ella no constituye un deber, por lo que –a su vez– no puede considerarse inmoral. Los actos que causan un daño a otro son vicios morales, mientras que las faltas autorreferentes no son propiamente inmorales.<sup>23</sup>

Mill establecía el límite de la moral –entonces– en el daño a un tercero porque creía que la sociedad tendía a castigar cualquier conducta considerada inmoral: si la moral comprendía deberes para con uno mismo, la tendencia de la sociedad se inclinaría a sancionar legalmente su incumplimiento, con consecuencias nefastas para la autonomía y privacidad de los ciudadanos.

Pero esta –desde luego– no es la única alternativa disponible. Podemos sostener a la vez: a) que la moral comprende deberes para con uno mismo, y b) que la sociedad solo puede intervenir mediante el castigo en casos de daño causado a un tercero sin su consentimiento. En esta situación, que representa la alternativa que yo elegiría, deberíamos aceptar que la RO representa un resumen incompleto de la moral.

## Motivación y auto-evidencia

La RO no menciona ningún motivo para obedecer a su prescripción. ¿Es ella misma una regla automotivadora? Algunos autores<sup>24</sup> sostienen que el hecho de invocar esta regla muestra en la sociedad un sentimiento altruista, puesto que la RO constituye un llamado a la generosidad en las interacciones sociales humanas. Este es un error. La RO no invoca la generosidad, puesto que prescribe que un egoísta trate de manera egoísta a los demás; lo que la RO invoca –en cambio– es un sentimiento de reciprocidad.<sup>25</sup>

La RO parecería ser automotivadora en el sentido de que apela a nuestros deseos sin filtrarlos, esto es, sin decir qué tipo de deseos deben ser tomados en cuenta por la regla: la RO no semeja decirnos lo que deberíamos desear, sino, en cambio, que los deseos y quererles deben convertirse en el criterio para la moral.<sup>26</sup> No distingue, por ejemplo, entre deseos altruistas y deseos autorreferentes. Esta es

<sup>23</sup> JOHN STUART MILL, *An Examination Of Sir William Hamilton's Philosophy*, University of Toronto Press, 1979, capítulo XXVI, p. 454; "On Liberty", en *Three Essays*, Oxford University Press, 1975, p. 93.

<sup>24</sup> Cfr. CHRISTOPHER BOEHM, "How the Golden Rule can Lead to Reproductive Success", en el Simposio cit., pp. 2 y 31.

<sup>25</sup> Boehm reconoce luego esto, y dice que la cooperación requiere reciprocidad, luego de lo cual distingue diferentes tipos de reciprocidad. BOEHM, cit., pp. 8-9.

<sup>26</sup> WATTLES, *The Golden Rule*, cit., pp. 5 y 7.

una interpretación parcial, sin embargo. El punto de partida de la RO es humeano, pero la RO dista de ser completamente humeana: yo deseo que me traten de una cierta manera, es cierto, pero no tengo por qué desear también que yo trate a los demás como deseo –de hecho– que ellos me traten a mí.

Si la RO es un concepto, tal vez por eso mismo se la considera autoevidente; los conceptos suelen ser autoevidentes, por cierto. “Dar a cada uno lo suyo” – como vimos– es un concepto de justicia y es –desde luego– autoevidente: nadie discutiría con otra persona por la posesión de un bien diciendo “Dame lo que es tuyo”. La discusión comienza cuando se ha formulado una concepción y comienza a determinarse qué es lo suyo de cada uno.

Pero la RO, sin especificaciones ni agregados, no es autoevidente en la forma en que sí lo es la idea de dar a cada uno lo suyo; los numerosos contraejemplos a la RO son una muestra clara de su falta de autoevidencia (y de sus diferencias, entonces, con la idea ortodoxa de un concepto). ¿Se requiere entonces alguna razón para obedecer la RO? Es lo que Locke piensa en el *Ensayo sobre el Entendimiento Humano*: no existe ninguna regla moral que se haya propuesto respecto de la cual un hombre no pueda justamente requerir una razón.<sup>27</sup> Y agrega que si fuera sugerida la regla “Uno debería hacer lo que querría que le hicieran a uno” a una persona que nunca la escuchó antes, ella podría preguntar –sin caer en el absurdo– cuál es la razón por la que debería seguir esa regla.

Si estamos buscando un axioma que posea la propiedad de ser autoevidente, la RO no sirve entonces como axioma, porque no cumple con el requisito de la intuitividad. Las bondades universales de la reciprocidad no son autoevidentes, como es autoevidente –supongo, al menos– lo bueno de perseguir nuestra felicidad. Voy a sugerir luego que la RO debe emplearse como complemento al principio general utilitarista, cumpliendo una tarea necesaria para que el utilitarismo no se confunda con el egoísmo: es un axioma, sí, pero no un axioma autoevidente.

## La crítica kantiana

En una nota de pie de página que aparece en la *Fundamentación de la Metafísica de las Costumbres*, Kant muestra una actitud decididamente crítica frente a la RO. La nota aparece dentro de un contexto en el cual Kant busca precisar la fórmula de ser un fin en sí mismo, ilustrándola con el ejemplo de los ataques a la libertad y a

---

<sup>27</sup> LOCKE, *Ensayo sobre el Entendimiento Humano*, libro I, capítulo 2, sección 4.

la propiedad ajenas: el violador de los derechos del hombre intenta usar la persona de los demás como un medio, sin tomar en consideración que ellos –en tanto seres racionales– deberían ser considerados como fines. El texto de la nota es el siguiente:

*Que no se piense que la trivial Quod tibi non vis fieri, etc. puede aquí servir como un criterio o principio. Porque él es simplemente derivado de nuestro principio, aunque con varias limitaciones. No puede ser una ley universal, porque no contiene la base ni de los deberes para con uno mismo, ni de los deberes de amor respecto de los demás (porque muchos hombres alegremente consentirían que otros no los beneficiaran, si pudieran ser exceptuados de beneficiarlos a ellos). Ni, finalmente, contiene la base de los deberes estrictos hacia los otros, porque el delincuente podría sobre esta base ser capaz de discutir con los jueces que lo están castigando; y así sucesivamente.<sup>28</sup>*

Esta crítica es correcta en varios aspectos. Ante todo, no le asigno importancia al hecho de que Kant emplee la formulación negativa de la RO, porque las mismas objeciones que formula pueden aplicarse también a la versión positiva. Hay tres críticas centrales que aparecen expresadas aquí. a) La RO no se refiere a los deberes para con uno mismo. Esta observación es correcta y ya la he mencionado antes. Si, como Mill piensa, la moral no incluye deberes para con uno mismo, la observación no es una crítica, pero Kant sí cree que existen tales deberes. b) La RO no prescribe deberes de amor para con los demás. Esto es cierto en la versión que emplea Kant, aunque no lo es en la mayoría de las versiones religiosas, como hemos visto, las que se refieren –precisamente– al amor al prójimo. c) La RO permite al delincuente cuestionar la sentencia del juez. Este es un error de Kant, que ya he mencionado con anterioridad. Para mostrarlo con más claridad, nada mejor que recurrir al comentario de Hoche:

*Si el acusado fuera a decir ante el Tribunal: “A usted tampoco le gustaría ir a la cárcel, si hubiera robado un auto y finalmente tuviera un poco de dinero”, el juez podría concederle el punto y decirle: “Usted está completamente en lo correcto: en su lugar yo querría precisamente lo que usted quiere ahora, a saber, ser dejado solo para disfrutar mi botín en paz.*

<sup>28</sup> IMMANUEL KANT, *Groundwork of the Metaphysics of Morals*, New York, Harper & Row, 1964, capítulo II, p. 97.

Pero

*ese no es el punto. El punto es cómo yo deseo ahora ser tratado en ese caso hipotético, y en relación con esto solo puedo decir: de manera justa y de acuerdo a la ley”.*<sup>29</sup>

Sin embargo, el motivo del rechazo kantiano a la RO es mucho más profundo. Las diferencias de la RO con el imperativo categórico se centran en dos aspectos fundamentales: a) el imperativo categórico se concentra en principios de la voluntad, y b) el imperativo categórico se concentra en reglas para la acción y en un criterio racional para evaluarlas.<sup>30</sup> Frente a ello, la RO reemplaza los principios de la voluntad por algo que es anatema para Kant, el deseo,<sup>31</sup> y carece de todo criterio racional para evaluar los resultados de su aplicación.

Para empeorar las cosas, algún lector desprevenido podría sentir la tentación de identificar la RO con el imperativo categórico, y esto resultaba intolerable para Kant, porque la RO es antagonica a la filosofía moral kantiana.

## La defensa utilitarista

Pero muchas de las razones por las que Kant rechaza la RO son –justamente– razones para que Mill la acepte. Las más importantes, desde luego, son: a) el rol

---

<sup>29</sup> HOCHE, cit., p. 78. Creo que Judith Thomson comete el mismo error que Kant al replantear el conocido caso moral del tranvía. En la presentación usual del caso, un tranvía está fuera de control y se dirige hacia cinco personas atrapadas en una vía, a las que va a matar. Existe una alternativa, sin embargo: accionando una palanca –que está al alcance de un paseante– el tranvía puede cambiar de dirección y dirigirse a otra vía, donde solo hay atrapada una persona. ¿Debe el paseante accionar la palanca? La mayoría de las personas interrogadas responde que sí, pero Thomson agrega ahora una tercera opción: también se puede cambiar la dirección del tranvía accionando la palanca para el otro lado, solo que en este caso matará al paseante. Thomson introduce luego subrepticamente la RO, recuerda que el paseante no desea morir, y de aquí infiere que tampoco puede entonces matar al individuo que está solo en la vía, puesto que no desea que esto le pase a él. (JUDITH JARVIS THOMSON, “Turning the Trolley”, *Philosophy and Public Affairs*, vol.36, number 4, pp. 361 y 364-365) Esto es un error. Por supuesto que si yo fuera el paseante no querría morir, pero lo que interesa es otra cosa: ¿qué regla querría yo que se aplicara en casos hipotéticos semejantes, en los que tal vez podría estar involucrado como víctima potencial? Y la respuesta es muy simple: aquella regla que minimice el número de víctimas. Lo que cuenta, entonces, no es mi reacción frente al caso en el que estoy involucrado, sino mi reacción frente al caso hipotético. En resumen, sigamos desviando el tranvía.

<sup>30</sup> WATTLES, *The Golden Rule*, cit., p. 84.

<sup>31</sup> Claramente el imperativo categórico prohíbe toda referencia a los deseos. Cfr. DANIEL BERTHOLD, “The Golden Rule in Kant and Utilitarianism”, en el Simposio cit., p. 13.

que desempeña el deseo en la RO, el cual –a diferencia de lo que ocurre con la teoría kantiana– puede ajustarse a las pretensiones del utilitarismo; y b) el alcance que la RO le asigna a la moral, cuyos límites coinciden con las conductas que afectan a terceros, exactamente como hemos visto que sostiene Mill.

Es por eso que no sorprende demasiado encontrar esta afirmación en su *Utilitarismo*:

*Entre su propia felicidad y la de los otros, el utilitarismo requiere [del agente] que sea tan estrictamente imparcial como un espectador benevolente y desinteresado. En la regla de oro de Jesús de Nazareth, leemos el espíritu completo de la ética de la utilidad. Hacer como uno querría que le hicieran y amar al prójimo como a nosotros mismos constituyen la perfección ideal de la moral utilitarista.<sup>32</sup>*

Mill utiliza tanto la formulación moral (esta vez, a diferencia de Kant, en su versión positiva), como la formulación religiosa. Pese a su enfática afirmación, empero, es difícil reconciliar al utilitarismo con el cristianismo, puesto que nadie sostendría seriamente que Jesús concedía importancia moral a las consecuencias de las acciones. Lo más que podría afirmarse es que el utilitarismo y el cristianismo pueden llegar –y solo en algunos casos– a resultados similares, aunque partiendo de fundamentos diferentes. Como observa Gensler,<sup>33</sup> los utilitaristas podrían ver a la RO como un principio útil para promover el bien de la sociedad: los grupos que siguen a la RO prosperarían más que aquellos que no lo hacen. Este resultado, no obstante, atraería la aprobación de un evolucionista, más que de un cristiano.<sup>34</sup> De todas formas, Mill necesita inevitablemente recurrir a la RO por una fuerte razón que mostraré en la sección que sigue.

Otro utilitarista que mira con simpatía la RO es Sidgwick, quien la consideraba un razonamiento moral usual. Él sostiene que

*cualquier acción que cualquiera de nosotros juzgue que es correcta para uno mismo, implícitamente juzga que es correcta para todas las personas similares en circunstancias similares. De aquí que pueda formularse una proposición respecto de lo que puede ser hecho a –no*

<sup>32</sup> JOHN STUART MILL, *Utilitarianism*, Indianapolis, The Bobbs-Merrill Co., 1971, capítulo 2, p. 24.

<sup>33</sup> GENSLER, *Formal Ethics*, cit., p. 134.

<sup>34</sup> Cfr. BERTHOLD, cit., p. 20.

*por— diferentes individuos (...). Así aplicada, ella aparece en lo que es popularmente conocido como la Regla de Oro: “Haz a los demás como querrías que ellos te hicieran a ti”.*

Sidgwick está anticipando el análisis que más tarde haría Hare en Freedom and Reason. Está vinculando primero la moral con la universalidad y, luego, la universalidad con la RO. En un comienzo, como estamos viendo, Sidgwick emplea la versión positiva de la RO, vinculándola con una acción, pero enseguida advierte que esto no es conveniente, puesto que

*esta fórmula es obviamente imprecisa en su enunciado; porque uno podría desear para otro la cooperación en el pecado, y estar dispuesto a reciprocársela.<sup>35</sup>*

Por eso Sidgwick piensa que el principio

*debe adoptar alguna forma negativa, como esta: “No puede ser correcto para A tratar a B de una manera en la cual sería incorrecto que B tratara a A, simplemente sobre la base de que ellos son dos individuos diferentes”.*

Pero no debe confundirse esta versión definitiva de Sidgwick con una formulación negativa de la RO. La auténtica formulación negativa requiere que se postule una omisión, y Sidgwick no lo hace: no dice “No hagamos a los demás...”, como ordena la versión negativa de la RO. Sidgwick introduce la negación en otra parte, en lugar de decir “Es correcto que hagas a los demás lo que quieres que te hagan a ti”, dice “No es correcto que hagas a los demás lo que no quieres que te hagan a ti”, pero creo que esta es una versión positiva de la RO, porque sigue requiriendo alguna acción y no una omisión. Por supuesto que Sidgwick —a diferencia de Mill— aceptaba que la RO no podía proporcionar una versión completa de la moral.<sup>36</sup>

Los utilitaristas, entonces, se muestran partidarios de la RO, en marcada diferencia al kantismo, y por eso es hora de ocuparnos de la forma que debería adoptar la RO para satisfacer los requerimientos de la teoría utilitarista. Pero antes de hacerlo,

---

<sup>35</sup> Sidgwick considera aquí el caso en el cual una de las partes tiene deseos inmorales, pero ignora el caso —de igual importancia— en el cual las dos partes tienen deseos inmorales.

<sup>36</sup> HENRY SIDGWICK, *The Methods of Ethics*, London, MacMillan, 1962, pp. 379-380.



atendamos a una advertencia imprescindible. El utilitarismo se propone dos cosas: a) proporcionar un criterio para identificar lo bueno, y b) proporcionar un criterio para identificar lo que se debe moralmente hacer. La tarea descrita en a) es propia de la ética axiológica, que se ocupa de lo bueno, mientras que la tarea descrita en b) es propia de la ética normativa, que se ocupa de lo debido. Recordemos –entonces– que el utilitarismo pretende ser tanto una ética axiológica, como una ética normativa.

## La Regla de Oro y la teoría utilitarista

La RO exhibe el imperativo de que hagamos a los demás lo que querríamos que nos hicieran a nosotros; puesto que el agente utilitarista reclama que se reconozcan sus deseos, la RO le exige que él reconozca también los deseos de los demás. ¿Cómo podría formularse, entonces, la RO desde una perspectiva utilitarista?

Atendamos primero a la distinción entre el utilitarismo de estados mentales y el utilitarismo de satisfacción de preferencias. En el primer caso la RO diría: “Trata a los demás de modo de causarles la mayor felicidad (de un modo compatible con la mayor felicidad general), tal como querrías que ellos lo hicieran contigo”. En el segundo caso la RO diría: “Trata a los demás de modo de satisfacer máximamente sus preferencias (de un modo compatible con la satisfacción general de las preferencias), tal como querrías que ellos lo hicieran contigo”.

Ambas versiones de la RO incorporan una asimetría que evita las críticas más frecuentes, tal como la crítica de Bernard Shaw que ya he mencionado: esta asimetría muestra, por ejemplo, que no debo necesariamente castigar a los demás agentes solo porque a mí me guste ser castigado, porque es posible que a ellos el castigo no los haga felices, o que no satisfaga sus preferencias.

Prestemos atención ahora a la distinción entre omitir y actuar: las dos formulaciones propuestas exigen actuar, de modo que si se limitaran a exigir omisiones dirían otra cosa. En el caso del utilitarismo de estados mentales, la RO diría “No trates a los demás de modo de causarles infelicidad (salvo que ese trato maximice la felicidad general), como no querrías que ellos lo hicieran contigo”. En el caso del utilitarismo de satisfacción de preferencias, la RO diría “No frustres las preferencias de los demás (salvo que esa frustración maximice la satisfacción general de las preferencias), como no querrías que ellos lo hicieran contigo”.

Yo creo que la teoría utilitarista es una teoría exigente, y que impone al agente acciones, no omisiones, de donde estas dos últimas versiones no alcanzarían

a satisfacer los recaudos de la teoría. No debe confundirnos en este aspecto la existencia de un utilitarismo negativo, puesto que el utilitarismo negativo exige acciones, tanto como lo hace el positivo: la diferencia entre ambos consiste en que el utilitarismo negativo exige acciones para disminuir el dolor y el utilitarismo positivo —en cambio— exige acciones para aumentar la felicidad; pero ambos, como se ve, exigen acciones.

Consideremos, por último, otra diferencia posible. En todas las versiones formuladas he considerado el estado mental de felicidad y la satisfacción de las preferencias, sin cualificación ninguna. Sin embargo, esta actitud es admisible solo si analizamos la versión utilitarista de Bentham, pero no la de Mill. Para Bentham, los estados mentales se ordenan teniendo en cuenta solo la cantidad de felicidad que reside en ellos, sin importar su causa. Para Mill, en cambio, debemos tender a producir los placeres superiores, y no los inferiores. La RO milliana diría: “Trata a los demás de modo de causarles la mayor felicidad derivada de los placeres superiores (de un modo compatible con la felicidad general), tal como querrías que ellos lo hicieran contigo”, o “Trata a los demás de modo de satisfacer máximamente sus preferencias por los placeres superiores (de un modo compatible con la satisfacción general de las preferencias por los placeres superiores), tal como querrías que ellos lo hicieran contigo”. Esta diferencia es importante, puesto que ella refleja la diferencia entre deseos e intereses: la RO de Bentham es liberal, en tanto atiende a los deseos del agente, mientras que la RO de Mill es paternalista, puesto que atiende a los intereses del agente.

Todas las versiones que he mencionado aquí comparten este rasgo, que es fundamental tener en cuenta: la RO —en todas ellas— es un precepto de ética normativa, y no de ética axiológica; no dice lo que es bueno, sino lo que se debe hacer, y este es un rasgo cuya importancia mostraré enseguida.

El problema que afecta a todas estas formulaciones es que el utilitarismo no parece necesitar la cláusula final (“tal como querrías” o “como no querrías”) de ninguna de ellas, porque la orden que imparte el utilitarismo se encuentra en la cláusula inicial, y esto es todo lo que la teoría necesita. ¿Qué papel cumple entonces esa cláusula final, que es la que convierte la máxima ortodoxa del utilitarismo en una RO?

¿Es acaso una justificación de la cláusula inicial? Para serlo, la máxima utilitarista no debería tener el papel de axioma sino de teorema, y la RO pasaría a ser el único axioma, papel que le resulta inadecuado, porque la máxima original tiene para todos los utilitaristas el carácter de regla suprema, y para muchos de ellos también el carácter de regla autoevidente, carácter que ya hemos visto que la RO no posee.

Podría tener, en cambio, el carácter de una explicación: “No trates a los demás... porque no querías...”. Puedo coincidir con esta idea, pero me parece que el papel central de la RO en la teoría utilitarista es otro, y está vinculado con el paso de la ética axiológica a la ética normativa, como enseguida veremos. Por otra parte, para Bentham y para Mill el problema central era diferente, y consistía en distinguir el utilitarismo del egoísmo, esto es, el problema de pasar de desear la felicidad individual a desear la felicidad general. Ninguno de ellos resolvió el problema convincentemente, pero creo que Hare proporcionó una solución satisfactoria, que se completa luego con la RO.

El problema de pasar de la felicidad individual a la general es crucial para el utilitarismo, pues –como dije– esto es lo que lo diferencia del egoísmo ético. (Comencé el examen de este tema en el capítulo 7). Bentham vacila cuando enfrenta la cuestión: algunas veces parece creer que la simpatía ayudaría a resolverlo,<sup>37</sup> y otras veces la crítica con vehemencia.<sup>38</sup> Mill reconoce que hay una diferencia crucial entre la felicidad individual y la general, pero se limita e recurrir a un supuesto sentido de conciencia que poseería la humanidad, el cual –por su parte– produce el deseo de encontrarse en unidad con las restantes criaturas.<sup>39</sup>

Desde luego que todos estos argumentos son altamente insatisfactorios, y me parece que Mill –conciente de ello– piensa que la RO es el puente que le permite pasar de la felicidad individual a la general. El utilitarismo parte –desde luego– de la felicidad individual: lo bueno es lo que aumenta la felicidad del agente,<sup>40</sup> y esa felicidad es lo único deseado en la realidad.<sup>41</sup> Hasta este momento, la teoría utilitarista presenta dos rasgos: a) es una teoría motivadora, porque claramente no hay dificultad alguna en concebir que todos los agentes desean su propia felicidad, y b) es una teoría indistinguible del egoísmo ético. (Es hasta ahora –además– solo una teoría de ética axiológica).

El principio básico utilitarista dice que lo bueno es la felicidad, y la teoría nos dice –de un modo verosímil– que todos deseamos nuestra felicidad. Supongamos ahora que incorporamos la RO: debemos desear para los demás lo que deseamos para nosotros. Con estos dos principios, podemos construir este argumento:

<sup>37</sup> JEREMY BENTHAM, “Tables of the Spring of Actions”, en *The Works of Jeremy Bentham*, Bristol, Thoemmes Press, 1995, vol.I, p. 212.

<sup>38</sup> JEREMY BENTHAM, *An Introduction to the Principles of Morals and Legislation*, Oxford, Basil Blackwell, 1967, capítulo II, sección 12.

<sup>39</sup> MILL, *Utilitarianism*, cit., pp. 33-34.

<sup>40</sup> BENTHAM, *Introduction...*, cit., p. 126.

<sup>41</sup> MILL, *Utilitarianism*, cit., p. 39.

- a) Deseo mi propia felicidad (y esto es –en consecuencia– lo bueno).
- b) Debo desear para los demás lo que deseo para mí.
- c) Debo desear la felicidad de los demás.<sup>42</sup>

La conclusión c) indica que debo desear la felicidad general, que es lo que los utilitaristas quieren demostrar en términos de ética normativa. No puedo llegar a esta conclusión solo con la premisa a), que es el principio básico utilitarista, pero puedo hacerlo si complemento la premisa a) con la premisa b) que es –precisamente– la RO. El principio básico utilitarista más la RO conducen a la conclusión utilitarista. (Quiero advertir –no obstante– que no es este el papel definitivo que le asigno a la RO, sino el que muestro en la sección siguiente).

Nótese que no se puede pasar directamente de la premisa a) a la conclusión c) –entre otras cosas– por un motivo muy sencillo: a) es una proposición de la ética axiológica, mientras que c) es una proposición de la ética normativa; se requiere entonces la introducción de b), que es también una proposición de la ética normativa.

Creo que Mill advirtió esto y por eso elogió de un modo tan caluroso la RO. Creo también que este es un modo de interpretar la teoría de Hare en *Freedom and Reason*, donde la exigencia de la universalidad se asemeja a la introducción de la RO, aunque enseguida veremos que la universalidad –al menos en una de sus acepciones– cumple un papel distinto al de la RO.

Queda ahora, sin embargo, un problema pendiente: ¿cuál es el *status* de la RO dentro de la teoría utilitarista? No puede ser un teorema, porque no veo cómo ella podría inferirse del axioma contenido en la premisa a). Me parece que no cabe otra respuesta que aceptar que se trata de otro axioma del utilitarismo, un axioma más incómodo que el primero, por supuesto, porque no es autoevidente. Pero es un axioma, sin duda, porque puede verse que a) no deriva del primero sino que es independiente de él, y b) resulta necesario para alcanzar la conclusión central de la teoría.

Que el utilitarismo tenga –hasta ahora– dos axiomas en vez de uno, como usualmente se lo entiende, no significa que la teoría pase a ser dualista. La teoría sigue siendo monista en el sentido que importa en ética, esto es, en el sentido de sostener que existe un solo valor, en este caso, la felicidad general.

---

<sup>42</sup> La RO puede entenderse en términos de querer o de deber. En el primer caso dice: "Si quieres que te traten de la manera x, quiere también tratar a los demás de la manera x". En el segundo caso dice: "Debes tratar a los demás como quieres que te traten a ti". No voy a explorar aquí esta eventual diferencia, puesto que me interesa la RO desde el punto de vista de la ética normativa, de donde la entenderé en términos de deber. Por otra parte, no es difícil inferir el deber a partir del querer.

## El rol de la Regla de Oro en la teoría utilitarista

¿Qué conclusiones podemos extraer de todo esto? Voy a comenzar con una tibia defensa de la RO en su formulación usual: “Haz a los demás lo que quieras que te hagan a ti”. Supongamos que yo soy un masoquista y quiero que me produzcan dolor. Parece seguirse entonces de la RO que yo debo causar dolor a los demás, pero esta es, por cierto, una lectura muy superficial. No se requiere un gran esfuerzo por parte del masoquista para entender que lo que él desea es que se satisfagan sus preferencias (masoquistas), de donde tampoco requiere un gran esfuerzo de su parte entender que la RO le pide que satisfaga también las preferencias (no masoquistas) de los demás, las que no incluyen sentir dolor. Entender la RO de otra manera conlleva una cierta dosis de mala fe. En realidad, es difícil que se conciba la RO como la única guía para la elección práctica, independientemente de todos los otros códigos de conducta.<sup>43</sup> Yo la empleo, por ejemplo, unida a los preceptos utilitaristas.

Continúo ahora con la formulación religiosa “Ama al prójimo como a ti mismo”. Si el agente se odia a sí mismo, dice el crítico, la RO le ordena odiar al prójimo. Esto también conlleva una cierta dosis de mala fe, puesto que la RO religiosa debe entenderse como un condicional: “Si te amas a ti mismo, ama al prójimo como a ti mismo”. De todas formas, y para no correr riesgos, el cristianismo ordena amar al prójimo como Jesús ama a su pueblo.

¿Y qué ocurre en el ámbito filosófico? Es inútil buscar partidarios de la RO dentro de la filosofía kantiana, porque la RO es la antítesis del kantismo: la RO recurre a los deseos y la filosofía kantiana prescinde de ellos. Para colmo, algún lector desprevenido puede estar tentado de confundir la RO con el imperativo categórico, y el kantismo debe poner fin desde el comienzo a esta herejía y repudiar la RO.

Pero es muy distinto el caso del utilitarismo. Supongamos que yo valoro únicamente mi felicidad individual, esto es, supongamos que –hasta ahora– soy un egoísta. Aquí enfrento dos alternativas:

- a) Yo quiero hacer todo lo que resulte en un incremento de mi felicidad individual, y quiero –además– que todo el resto de la gente haga todo lo que resulte en un incremento de *mi* felicidad individual: esta es mi idea de lo bueno. Una actitud tal puede ser racional, pero –ciertamente– no es razonable. Amén de ello, no es una postura ética, puesto que no respeta el requisito de la universalidad: yo soy hasta aquí –solamente– un egoísta psicológico.

<sup>43</sup> Cfr. TED HONDERICH (ed.), *The Oxford Companion to Philosophy*, Oxford University Press, 1993, p. 321.

b) Yo quiero hacer todo lo que resulte en un incremento de mi felicidad individual, y acepto que el resto de la gente haga todo lo que resulte en un incremento de su felicidad individual. Esta es mi idea de lo bueno. Una actitud tal es –a la vez– racional y razonable, a la par de ser una postura ética: yo soy ahora un egoísta ético y no solamente un egoísta psicológico, porque he incorporado el requisito de la universabilidad de mis creencias.

Agreguemos ahora la RO, para arribar a c): puesto que yo quiero hacer todo lo que resulte en un incremento de mi felicidad individual, y me gustaría que los demás contribuyeran a esto, también debo hacer todo lo que resulte en el incremento de la felicidad de los demás. Ahora, a la postura inicial egoísta psicológica descrita en a), no le incorporo únicamente la acepción de la universabilidad mencionada en b), porque ella solo me conduciría al egoísmo ético, sino que le he agregado la RO, la cual cumple entonces dos funciones: arribar al utilitarismo y obtener una teoría de ética normativa. (Puede cuestionarse la primera función y sostenerse que la universabilidad por sí misma –en alguna otra acepción– permite arribar directamente al utilitarismo, sin intermediarios, pero lo que no puede discutirse es que –en este caso– se trataría solo de un utilitarismo axiológico, y no normativo).

Para explicar nuevamente la forma en que funciona la universabilidad, y su alcance, voy a formular un argumento de tres pasos. Primero, si una persona dice que una cosa es colorada, está comprometida a decir que cualquier otra cosa que sea como ella en sus aspectos relevantes será igualmente colorada; y los aspectos relevantes son –precisamente– aquellos que la autorizan a llamar colorada a esa cosa. Este es un juicio descriptivo. Segundo, los juicios morales tienen también un significado descriptivo, aunque en este caso las reglas que determinan este significado son principios morales sustantivos. Tercero, siendo esto así, si una persona dice que la felicidad individual es buena, está comprometida a decir que la felicidad de las demás personas es buena.<sup>44</sup>

Por eso los utilitaristas no solo simpatizan de hecho con la RO: necesitan de la RO, al menos, para pasar de la ética axiológica a la normativa (así como también la necesitan para pasar de la felicidad individual a la general). La necesitan como un axioma, que complementa el axioma inicial de que lo único bueno es la felicidad del agente, y el axioma de la universabilidad. Nótese que el utilitarismo no puede generarse solo con el axioma de que la felicidad del agente es lo bueno, y nótese también que la universabilidad no puede inferirse del primer axioma como un

---

<sup>44</sup> Los dos primeros pasos del argumento, por supuesto, son de HARE, cit., pp. 11, 15 y 30. Creo que él podría discrepar con el alcance del argumento.

teorema, de donde la universalidad es –en realidad– el segundo axioma de la teoría.<sup>45</sup> El tercer axioma (la RO) no es autoevidente, como sí lo es el primero, y tampoco –entonces– motiva, como sí motiva el primero. Pero tampoco es autoevidente y tampoco motiva el axioma de la universalidad. La teoría completa no es internalista, como sí lo es claramente el egoísmo ético, salvo que se presuma en los agentes morales un sentido innato de simpatía.

En conclusión, el utilitarismo requiere tres axiomas: a) lo bueno es la felicidad del agente; b) si la felicidad individual es buena, la felicidad de las demás personas es buena; y c) la RO. Los dos primeros axiomas pueden ser suficientes para generar un utilitarismo axiológico, pero el utilitarismo tiene pretensiones de ser también una ética normativa.

Lo que no me parece correcto, en cambio, es la identificación milliana del cristianismo con el utilitarismo. El cristianismo, como ya he advertido, propone una moral deontológica, despreocupada de las consecuencias, esto es, una posición moral opuesta a la del utilitarismo, donde son las consecuencias las que determinan la corrección o incorrección de las acciones. Tal vez alguien querría sostener que el cristianismo y el utilitarismo tienen distintos fundamentos, pero llegan –sin embargo– a los mismos resultados. Esta es una identificación más débil –desde luego– pero todavía la encuentro exagerada; creo que el cristianismo y el utilitarismo tienen diferentes fundamentos y que –en muchos casos cruciales– también llegan a diferentes resultados.

Pienso que he explicado –entonces– cuál es el papel que cumple la RO en la filosofía moral. Como sin duda ese papel lo tiene solo dentro de la teoría utilitarista, me pareció adecuado hacerlo notar desde el título del capítulo, y hablar así de la regla de oro del utilitarismo.

---

<sup>45</sup> Hay un segundo sentido de universalidad en Hare que puede equivaler a la RO. Ocurre cuando él sostiene que "se sigue de la universalidad que si yo ahora digo que debo hacer una cierta cosa a una cierta persona, estoy comprometido con el punto de vista de que la misma cosa me debería ser hecha a mí si yo estuviera exactamente en su situación". De esta acepción se sigue correctamente que "el requerimiento de universalizar nuestras prescripciones genera el utilitarismo", en tanto ética normativa. Cfr. R. M. HARE, *Moral Thinking*, Oxford, Clarendon Press, 1981, pp. 108 y 111.

