

UTILITARISMO DE REGLAS Y DERECHOS LIBERALES*

JONATHAN RILEY
Tulane University

ABSTRACT

Liberal rule utilitarianism seems to provide a «secure and acceptable basis» for equal liberty, contrary to suggestions by Rawls and others. Harsanyi's variety gives absolute priority to some system of equal rights and liberties, by virtue of the fact that an optimal moral code (and its permissible set of strategies) has absolute priority over all nonmoral considerations. On the other hand, Millian liberal utilitarianism can even accommodate absolute liberal priorities *within* its optimal code, although modifications to Harsanyi's orthodox utility theory are required to do so.

SUMARIO: I. Introducción; II. El utilitarismo de reglas liberal de Harsanyi; III. Prioridades absolutas dentro de un código utilitarista liberal; IV. Conclusión.

* Traducción de Joaquín R.-Toubes. Versiones previas de este trabajo fueron presentadas en el Congreso de Caen sobre «Justicia, Liberalismo político y utilitarismo» celebrado en la Universidad de Caen, Francia, del 14 al 17 de junio de 1996; y en el Congreso de Lisboa sobre «Derechos y utilitarismo», celebrado en Arrabida, Portugal, del 11 al 15 de Junio de 1997. Aparecerá publicado en Maurice Salles y John Weymark, eds., *Justice, Political Liberalism, and Utilitarianism: Proceedings of the Caen Conference in Honour of John Harsanyi and John Rawls* (Cambridge: Cambridge University Press). Agradezco a Ken Binmore, Charles Blackorby, John Broome, David Donaldson, John Weymark y a un informante anónimo por sus útiles comentarios. La responsabilidad por las opiniones expresadas sigue siendo mía.

I. INTRODUCCIÓN

Es típico de los liberales afirmar que los derechos y libertades en igualdad deben tener una adecuada prioridad moral sobre los valores opuestos. Rawls (1971, 1993, 1995), por ejemplo, sostiene que es propio de la justicia social limitar las concepciones permisibles de la vida buena en una cultura democrática, y que, dentro de la justicia tal como él la concibe, un primer principio de libertades básicas iguales tiene prioridad absoluta sobre un segundo principio bipartito. Los derechos políticos ocupan un lugar especial dentro de su primer principio en tanto que ellos (a diferencia de otros derechos básicos) deben tener aproximadamente la misma valía (o «valor equitativo») para todas las personas (1993, pp. 5-6, 324-31, 356-62 [en la versión española: pp. 35-6, 361-8, 395-401]). La razón para este trato especial parece ser el papel supuestamente esencial que juega el ejercicio de las libertades políticas para «la conservación de las demás libertades» (ibid., p. 299 [336]). Así, los sujetos deben tener una oportunidad equitativa de ejercer sus derechos políticos básicos «porque resulta esencial a la hora de determinar una legislación justa y también para asegurar que el proceso político equitativo definido por la constitución está abierto a todos sobre la base de una igualdad aproximada» (ibid., p. 330 [368]).

Un objetivo primordial de los teóricos liberales en general, y de Rawls en particular, ha sido proporcionar a la adecuada prioridad de los derechos iguales un fundamento más sólido que el que parece ofrecer el utilitarismo. En palabras de Rawls, la «primera tarea» de su teoría política es «proporcionar una base más sólida y aceptable para los derechos básicos y los principios constitucionales que la que parece permitir el utilitarismo» (1985, p. 226). Implícitamente parece decirse que el utilitarismo es incapaz de proporcionar una garantía razonable del tipo exigido por la cultura política liberal, esto es, una garantía de que se privilegiará debidamente la igualdad de derechos y libertades sobre los valores opuestos.¹

Los utilitaristas modernos no han tardado en responder a ese desafío. Harsanyi (1977, 1992), por ejemplo, sostiene que el utilitarismo de reglas ofrece un fundamento razonable para las pretensiones liberales. En su opinión, los derechos están «moral-

¹ Para afirmaciones similares en ese sentido, véase también Rawls (1971), pp. 22-33, 315-7; y (1993), p. 37.

mente protegidos» en el sentido de que «*no pueden ser postergados simplemente porque al postergarlos incrementemos la utilidad social aquí y ahora, excepto posiblemente en algunos casos muy especiales donde están en juego intereses de la sociedad fundamentalmente importantes*» (1992, p. 689, cursivas originales). La igual libertad está asegurada por normas morales que todas las personas deben cumplir para maximizar la utilidad social a largo plazo, excepto quizás en casos aislados donde la suspensión de las normas está justificada para evitar una catástrofe social.

«Sólo el utilitarismo de reglas –insiste Harsanyi– puede explicar por qué una sociedad será mejor si la conducta social está limitada por una red de derechos morales y obligaciones morales que, salvo en emergencias extremas, no puede ser violada con base en consideraciones de mera eficiencia social» (1982, p. 41). Por el contrario, el utilitarismo de actos, que pide a cada persona escoger sin limitación la acción maximizadora de la utilidad social en cada situación, es «una moral super-maquiviélica» (ibid.). Los propios utilitaristas deben abandonarla, porque autoriza «la infracción de todos los derechos individuales y todas las obligaciones institucionales en nombre de cierta utilidad social definida de modo reducido», la cual es inferior a la que puede obtenerse con el utilitarismo de reglas (ibid.).

Sin embargo, pese a su perspectiva liberal, Harsanyi rechaza la idea de que la justicia social requiera reglas de prioridad absoluta. En particular critica a Rawls por sugerir que los conflictos «pueden resolverse por el procedimiento estrecho de miras de establecer rígidas *prioridades absolutas* entre los diferentes valores sociales, por ejemplo declarando que la *libertad* (o, más exactamente, la mayor libertad básica posible para todos que sea compatible con una libertad igual para todos los demás) tendrá prioridad absoluta sobre la solución de problemas de *desigualdad social y económica*» (1992, p. 696, cursivas originales). Es verdad –parece admitir– que un código moral utilitarista de reglas debe tener él mismo prioridad absoluta sobre otras consideraciones: «Ciertamente, hay un caso claro de *prioridad absoluta* de un valor social sobre otros. Es la prioridad absoluta que debemos atribuir a nuestros *deberes morales* sobre los intereses personales y sobre otras consideraciones no morales» (1996, p. 5, cursivas originales).² Pero el sentido común nos dice que las reglas rígidas de

² Pasaré por alto otras afirmaciones de Harsanyi (v.gr., 1985a, p. 54) que parecen contradecir ésta.

la prioridad absoluta deben ser evitadas *dentro* de un código óptimo:

El sentido común nos dice que la vida social está llena de situaciones donde tenemos que *sopesar* diferentes valores sociales unos con otros y debemos encontrar entre ellos *compromisos* moral y políticamente aceptables... Sin duda habrá casos donde el sentido común nos aconsejará aceptar una reducción *muy pequeña* de nuestras libertades si éste es el precio de una reducción *sustancial* en las desigualdades sociales y económicas (1992, p. 696, cursivas originales).

Se desprende de aquí que debemos aceptar que no se le otorgue a los derechos protección absoluta. Cualquier derecho concreto puede tener que ser sacrificado a veces en beneficio de derechos o deberes opuestos. La utilidad social puede exigir incluso que otros valores sociales, tales como una distribución equitativa de la riqueza, deban tener prioridad dentro del código sobre todas «nuestras libertades».

No obstante, la aversión de Harsanyi por las prioridades absolutas no es compartida por todos los utilitaristas. J.S. Mill (1859, 1861, capítulo 5), por ejemplo, sugiere que para maximizar la utilidad social en cualquier cultura avanzada ha de darse a algún conjunto de derechos y libertades iguales prioridad léxica sobre valores opuestos. Además, según Mill un subconjunto apropiado de derechos iguales debería prevalecer incluso frente a los demás derechos, aunque él difiere de Rawls respecto a la naturaleza del conjunto. Es fama que Mill afirma que la maximización de la utilidad social requiere derechos *absolutos* para escoger como a uno le plazca entre lo que él llamó acciones «puramente personales» [*self-regarding*, referidas a uno mismo] (1859, p. 224 [184]).³

Sostendré que un código utilitarista de reglas *debe* en principio dar prioridad absoluta a *algún* conjunto de derechos y libertades iguales sobre los valores sociales opuestos. Además, parece posible dar a algunos derechos prioridad absoluta sobre otros dentro de un código óptimo, aunque esto requiere modificar la teoría ortodoxa de la utilidad adoptada por Harsanyi.

³ Para más detalles sobre la doctrina de Mill sobre la libertad, véase Riley (1989-90), (1997a), (1997b).

II. EL UTILITARISMO DE REGLAS LIBERAL DE HARSANYI

Toda doctrina utilitarista supone que el peso asociado a los derechos iguales en relación a otras consideraciones debe ser determinado por la utilidad social. Por tanto, si ha de darse a la libertad igual un peso relativamente alto, los utilitaristas deben mostrar que un grado convenientemente alto de protección de un sistema de derechos iguales maximiza la utilidad social.

La modalidad de utilitarismo característica de Harsanyi ha evolucionado durante más de cuarenta años, como también lo ha hecho la versión del liberalismo característica de Rawls. Una vez advertido que pueden existir algunas tensiones entre sus diversas afirmaciones, puede interpretarse con justicia que Harsanyi ve los juicios morales utilitaristas como una especie (entre otras) de las preferencias racionales. Las preferencias son entendidas como jerarquizaciones consistentes de alternativas posibles. Tales jerarquizaciones no están basadas en, y no reflejan necesariamente, deseo-satisfacción o sentimientos de placer. Más bien se manifiestan en, y de hecho son equivalentes a, las conductas de elección racional que *serían observadas* si el agente se guiase por los axiomas definidores de una determinada concepción de la racionalidad.⁴ Además, como subraya Binmore (1994, p. 52), la teoría de la utilidad de Harsanyi es «completamente ortodoxa» entre los economistas neoclásicos. La utilidad es simplemente un término para el valor de una función que se emplea por comodidad para representar la conducta racional relevante: «Una función de utilidad, en el sentido moderno, no es más que un medio tratable matemáticamente de expresar el hecho de que la conducta electora de un individuo es *consistente*» (ibid., p. 51, cursivas originales).

Para ser llamadas utilitaristas, las preferencias de una persona deben satisfacer la compleja idea de Harsanyi acerca de un tipo moral de elección racional, en la cual están implicadas una enorme variedad de condiciones racionales, incluyendo: criterios

⁴ Es posible que la conducta electiva efectiva no revele preferencias racionales, al margen de cuál sea la concepción de la racionalidad. Por ejemplo, incluso aunque la racionalidad se agote en los axiomas de completud, reflexividad y transitividad que definen la idea de una ordenación, las elecciones efectivas podrían violar la transitividad, en cuyo caso el agente no se comporta como si tuviese una ordenación de preferencias. *A fortiori*, si se define la racionalidad de modo que incluya más condiciones, la conducta electiva efectiva podría violar esas condiciones, en cuyo caso el agente no se comporta como si tuviese una preferencia racional del tipo relevante.

de utilidad esperada; una exigencia de que para medir las intensidades de preferencias relativas deben usarse las actitudes hacia el riesgo y hacia la incertidumbre presentes en la forma de función de utilidad de von Neumann-Morgenstern (vNM); una condición de información plena; el criterio de equiprobabilidad, en virtud del cual un decisor moral debe olvidar sus actuales circunstancias y adoptar un punto de vista impersonal, imaginando que tiene la misma probabilidad de ocupar la posición social de cualquier persona con las preferencias de esa persona; la exigencia de Dworkin (1977, p. 234 [342]) de que las preferencias «orientadas a otros», configuradas respecto a las posiciones de otras personas, deban ser completamente ignoradas por los observadores impersonales; un postulado de similitud fundamental, que asegure que todos los observadores impersonales hagan elecciones morales y comparaciones interpersonales idénticas; y, finalmente, una exigencia de que los observadores impersonales se sometan conjuntamente a una autolimitación adoptando un compromiso vinculante hacia el mismo código moral óptimo.⁵ Cabe ver las condiciones de racionalidad relevantes como «imperativos hipotéticos» (en lugar de imperativos categóricos kantianos), en la medida en que su fuerza motivadora es contingente respecto a la posesión de ciertas actitudes, incluyendo actitudes hacia el riesgo y la incertidumbre así como actitudes hacia otras personas.⁶ Si sus preferencias satisfacen en efecto todas las condiciones relevantes, entonces la persona se comporta como si fuera un utilitarista de reglas (sean cuales sean sus deseos y sentimientos). Más aún, la utilidad social y moral es simplemente el nombre para el valor de una función (comparable cardinal e interpersonalmente), del tipo utilitarista tradicional, que representa las preferencias racionales de este tipo moral complejo.⁷

⁵ Para más detalles sobre la variedad de condiciones de Harsanyi, véase Riley (1997d), (1997e).

⁶ Puede ocurrir que la conducta de una persona satisfaga algunas, pero no todas, de las condiciones de racionalidad que comprenden el tipo moral de elección racional de Harsanyi. Así, es posible ser inmoral y no obstante ser racional en un sentido no moral más débil. Por ejemplo, Binmore (1994, 1997) no participa de la variedad moral de elección racional de Harsanyi, aunque comparte la idea de Harsanyi de que la utilidad representa simplemente la conducta racional, y también está de acuerdo con que una concepción bayesiana de la racionalidad es mejor que sus alternativas. En lugar de introducir axiomas apropiados (de elección impersonal y demás) para definir un tipo moral de elección racional, él mantiene separadas la moralidad y la racionalidad.

⁷ Broome sugiere que el uso que hace Harsanyi de «utilidad» «es ambiguo» porque «algunas veces usa el término ... como un nombre para el valor de una

Harsanyi sostiene que los utilitaristas tienen buenas razones para comprometerse con un código cuyas reglas permiten protección sustancial para la libertad igual. Dado que las personas tienen por naturaleza prejuicios a favor de sus intereses particulares en lugar de ser perfectamente imparciales, un código que distribuye derechos iguales para hacer algunas elecciones personales sin interferencias indebidas de otras personas, está más en consonancia con la naturaleza humana que un código como el del utilitarismo de actos, que no reconoce tales derechos. El utilitarismo de actos –insiste– tiene «efectos de aplicación negativos *intolerablemente gravosos*» (1992, p. 688, cursivas originales). Es simplemente demasiado exigente para los humanos, en comparación con dioses o santos. Sus «principios rígidamente universalistas» requerirían «una total supresión de nuestras inclinaciones racionales», las cuales son «particularistas» en el sentido de que el agente otorga «mayor peso» a los intereses propios, de su familia y de sus amigos que a los intereses de otras personas (ibid., p. 675, 688). Tal supresión «podría hacerse, si es posible hacerla, sólo con enormes esfuerzos y con costes psicológicos extremadamente altos» (ibid., p. 688). Así, un código utilitarista de reglas puede reconocer que la persona media gana utilidad si se le permite hacer sus propias elecciones al atender a sus intereses particulares, en lugar de forzarla siempre a cumplir su deber estrictamente imparcial tal como lo determina la maximización de la utilidad social en términos del utilitarismo de actos.⁸

Una segunda razón para escoger un código utilitarista de reglas que permite una protección sustancial de la igual libertad se refiere a los «efectos de creación de expectativas» positivos (entre ellos «efectos de incentivación» y «efectos de confianza») asociados con su sistema de derechos y deberes. Por ejemplo, las normas que

función que representa ... una ordenación de preferencias», pero otras veces lo usa «para significar bueno» (1991, p. 59, n. 19). Pero los dos usos no son necesariamente incompatibles. Para Harsanyi, la utilidad representa en efecto determinadas preferencias *racionales* definidas respecto a un conjunto relevante de objetos. Al mismo tiempo, los *criterios de racionalidad* relevantes también dan forma a las preferencias, las cuales constituyen el significado del bien. Así, la utilidad representa el bien y también la preferencia racional, precisamente porque el bien (incluidos el bien social y moral) es equiparado con la preferencia racional (incluidas social y moral).

⁸ Mediante la asignación de una «utilidad procesal» a la «elección moral libre», un código utilitarista de reglas podría también reconocer «la distinción tradicional, y muy atractiva intuitivamente, entre limitarse a cumplir el *deber* que uno tiene y realizar una acción *supererogatoria* que va más allá de la llamada del deber» (1992, p. 689, cursivas originales).

distribuyen y ponen en práctica derechos de propiedad privada sustanciosos, «proporcionan *incentivos* socialmente deseables para el trabajo duro, el ahorro, la inversión y las actividades empresariales» y «también dan a los titulares de propiedad *confianza* en cierta seguridad financiera y cierta independencia frente a la buena voluntad de otras personas» (ibid., p. 690, cursivas originales). En efecto, una «titularidad de propiedad extendida» genera «una suerte de efecto de confianza que beneficia a la sociedad en su conjunto», a saber, mayor «estabilidad social» y «libertad personal y política» (ibid.). El utilitarismo de actos no puede reconocer esos efectos positivos de creación de expectativas que tiene todo sistema completo de derechos y deberes iguales. Considera sólo los efectos de creación de expectativas de las acciones individuales, los cuales son «normalmente ... despreciablemente pequeños» (ibid.).⁹

Como admite Harsanyi, sin embargo, los efectos de incentiva-
ción y de confianza relevantes son realmente útiles socialmente sólo para las personas como nosotros, que estamos inclinadas por naturaleza a ser parciales en favor de nuestros intereses particularistas. A diferencia de los santos utilitaristas de actos, por ejemplo, puede que nosotros estemos poco dispuestos a participar en una empresa socialmente útil si no hay un sistema apropiado de derechos de propiedad privada y deberes correlativos. Los santos, en cambio, realizarían cualesquiera actos de trabajo y ahorro que fuesen necesarios para maximizar la utilidad social en términos utilitaristas de actos, con independencia de sus derechos a apropiarse de los frutos de su trabajo y ahorro. Similarmente, puede que nos neguemos a prestar dinero a personas relativamente pobres a menos de que tengamos la confianza de que se harán valer las obligaciones promisorias. A los santos, en cambio, no les importa si el sujeto pobre rompe su promesa de devolver una deuda, y «arrojarán más riqueza sobre él hasta que todos tengan la misma utilidad marginal por cada un más o menos» (Binmore, 1997, p. 157).¹⁰

⁹ Para un estudio ulterior de este importante punto, véase Harsanyi (1985a), pp. 42-6; y (1985b), pp. 70-72.

¹⁰ Binmore observa que no habría diferencia entre el utilitarismo de reglas y de actos en una sociedad ideal, donde todo el mundo es un utilitarista de actos «que participa de la misma visión profética de la naturaleza de la utilidad social» en condiciones de información perfecta, y donde todo el mundo siempre realiza conjuntamente aquellas acciones necesarias para maximizar la utilidad social (1997, p. 156, n. 24). Como Harsanyi, sin embargo, él considera el utilitarismo de actos demasiado exigente para los humanos. Como segunda mejor teoría «liberal

Se desprende que, para Harsanyi, un código utilitarista de reglas que privilegia apropiadamente los derechos iguales maximiza la utilidad social no tanto porque los seres humanos tienen deficiencias morales e intelectuales, sino porque el más racional y solidario entre nosotros, reconociendo nuestra naturaleza particularista, concedería un valor peculiar a la libertad del individuo para hacer sus propias elecciones, y a los incentivos y la confianza que hace posible una red razonable de iguales derechos y obligaciones. Las reglas utilitaristas liberales no son meramente herramientas para mitigar nuestras capacidades limitadas, las cuales el utilitarismo erróneamente cree ilimitadas. Son más bien restricciones socialmente útiles que conservan su utilidad para humanos, en cuando que distintos de los santos, incluso aunque supongamos que los humanos son genios completamente informados y perfectamente capaces de imaginarse en el lugar otros al tiempo que simulan tener los intereses particulares de esos otros.

Más formalmente, considérese el juego de n personas y dos etapas del utilitarismo de reglas con observancia plena en condiciones de información completa de Harsanyi (1992, pp. 692-4). La primera etapa es un juego cooperativo en el que todos los jugadores aceptan aplicar el modelo de equiprobabilidad para elegir un código óptimo M^* a partir de algún conjunto dado M de códigos morales practicables. Dado el postulado de similitud fundamental, podemos hablar de un solo observador impersonal que hace la elección en esta etapa. Cualquier código practicable $M \in M$ da lugar a un conjunto de estrategias permisible $P(M)$, que es el mismo para todos los jugadores que se vinculan a respetar el código.¹¹

La segunda etapa es un juego no-cooperativo en el cual cada jugador k es libre de elegir una estrategia personal (pura o mixta) s_k de entre su conjunto practicable de estrategias S_k de modo que maximice su utilidad personal, sometido al requisito de que:

$$\forall k: s_k \in P(M^*),$$

burguesa», él recomienda su propia alternativa humeana al utilitarismo de reglas (ibid., pp. 157-8). Para una discusión relacionada con este punto, véase Riley (1996).

¹¹ Harsanyi asume que, para todo $M \in M$, $P(M)$ es un subconjunto compacto no vacío del conjunto S de estrategias practicables, el cual por comodidad se supone igual para todo jugador, es decir, $S = S_1 = \dots = S_n$. Sin embargo, la última suposición podría fácilmente omitirse para permitir que cada jugador tenga un conjunto de estrategias diferente, en cuyo caso $S = \prod S_k$.

donde $P(M^*)$ es el conjunto de estrategias permisibles asociado con el código óptimo M^* elegido en la primera etapa. Cualesquiera derechos y obligaciones personales están incluidos entre sus estrategias permisibles. La idea de iguales derechos y libertades para todos se refleja en el hecho de que $P(M)$ es el mismo para todos. Cada persona tiene también una obligación general de abstenerse de hacer elecciones estratégicas no permitidas.

Para escoger un código óptimo M^* en la primera etapa, un observador impersonal debe comparar códigos practicables alternativos para determinar cuál de ellos está asociado con la mayor utilidad social. Los efectos de poner en práctica un código dado cualquiera están «representados por el hecho de que las estrategias de los jugadores estarán limitadas al conjunto permisible $P(M)$ definido por este código moral M » (ibid., p. 693). Los efectos de creación de expectativas están «representados por el hecho de que algunos jugadores escogerán diferentes estrategias que las que escogerían si su sociedad tuviera un código moral diferente —no porque M requiere directamente que lo hagan así, sino más bien porque esas estrategias son las *mejores respuestas* a las estrategias esperadas de los demás jugadores, sobre el supuesto de que esos *otros jugadores* sólo emplearán estrategias permitidas por el código moral M » (ibid., pp. 693-4, cursivas originales).

Pero para hacer una elección razonable, un observador ha de conocer cómo puede esperarse razonablemente que se comporten los jugadores en la segunda etapa, con cada uno de los códigos alternativos. Dado que todos observarán plenamente las reglas, se necesitan predicciones de las elecciones estratégicas permisibles que cada jugador hará efectivamente cuando interaccione con los demás, dando por supuesto que las elecciones de cada uno serán las mejores respuestas a las elecciones dadas de los otros. En pocas palabras, el observador impersonal debe ser capaz de predecir un punto de equilibrio de Nash $\bar{s} = (\bar{s}_1, \dots, \bar{s}_n)$ del juego no-cooperativo que se jugará en la segunda etapa. Aparecerán juegos no-cooperativos diferentes para códigos diferentes y sus respectivos conjuntos de estrategias permisibles. Dado un concepto de solución (véase, v.gr., Harsanyi y Selten, 1988), puede asumirse que el observador defina una «función de predicción» π que selecciona, para cada posible juego no-cooperativo $g(M)$, un equilibrio $\bar{s} = \pi(g(M))$.

Harsanyi supone que la función de utilidad personal u_k de cualquier jugador k toma la forma $u_k = u_k(\bar{s}, M)$. Él incluye M como una variable «porque los jugadores pueden derivar alguna utilidad directa por vivir en una sociedad cuyo código moral permite

una cantidad considerable de elección individual libre» (1992, p. 693). Sin embargo, puesto que la función de utilidad social y moral w de cualquier observador impersonal se define en términos de las funciones de utilidad personal u_1, \dots, u_n , debemos tener:

$$w = w(\bar{s}, M) = w(\pi(g(M), M).$$

La maximización de la utilidad social y moral se consigue mediante el compromiso conjunto hacia un código óptimo M^* que privilegia adecuadamente la libertad igual, de tal forma que cada persona k es libre de elegir una estrategia personal permisible $\bar{s}_k \in P(M^*)$ que es la mejor respuesta a las elecciones estratégicas permisibles de sus compañeros.

La estructura en dos etapas del juego pone de manifiesto que el observador impersonal está escogiendo un elemento óptimo entre un conjunto practicable de resultados *en equilibrio*. Ese conjunto practicable de combinaciones estratégicas en equilibrio es también el conjunto posible de sistemas de derechos y libertades iguales compatibles con la incentivación. En efecto, *previamente* a seleccionar un código óptimo, el observador debe resolver la cuestión de qué sistemas de derechos iguales son practicables, prediciendo los equilibrios de Nash relevantes. Así, hay un sentido en el que al valor de la libertad igual se le da prioridad absoluta sobre consideraciones distributivas y de otro tipo.¹²

Además, un código óptimo implica un sistema óptimo de derechos y libertades iguales, en el cual los pesos relativos asociados a las diferentes estrategias permisibles reflejarán cualquier proporción de consideraciones distributivas y otro tipo que el observador impersonal decida que maximiza la utilidad social. Si, por ejemplo, el observador decide que un deber de redistribuir la riqueza al necesitado posterga en algunas circunstancias el derecho a la propiedad privada, entonces el deber en cuestión es correlativo a un derecho superior del necesitado a la redistribución. Así, si un código moral óptimo tiene prioridad absoluta sobre consideraciones no morales, como aparentemente cree el propio Harsanyi, entonces un sistema óptimo de derechos y libertades iguales tiene prioridad absoluta sobre valores sociales opuestos.

Evidentemente, los jugadores en este juego utilitarista de reglas deben exhibir notables capacidades intelectuales y morales

¹² Para una discusión relacionada con este punto, véase Binmore (1994), pp. 9, 127-31, 147; y (1997), capítulo 3.

para identificar un código óptimo, incluso si aceptamos que su compromiso con tal código supone menores exigencias emocionales que el utilitarismo de actos. Aparte de imaginarse ellos mismos con la misma probabilidad en la posición de cada persona cuando hacen las elecciones particularistas que haría esa persona, esos agentes deben ser también capaces de predecir cómo cambiarán las interacciones de la gente a través de diferentes códigos y redes asociadas de derechos y obligaciones, para calcular qué código maximiza la utilidad social. Esas interacciones pueden ser extremadamente complicadas, pueden necesitar tener en cuenta sistemas de reglas altamente intrincados que implican muchos elementos excepcionales, reglas de orden superior para resolver contradicciones entre normas más básicas, etcétera.

Dadas las capacidades exigidas, los sujetos pueden maximizar la utilidad social comprometiéndose conjuntamente con un código óptimo M^* que privilegie adecuadamente la libertad igual. En virtud de ello lo justo y lo correcto se subsumen en la maximización de la utilidad social o moral; en otras palabras, en la concepción del bien del agente utilitarista de reglas. Su compromiso conjunto no sólo impide a esos sujetos rebajar la utilidad social desviándose del código óptimo para incrementar sus utilidades personales a expensas de sus compañeros. Además, como subraya Harsanyi, «el compromiso de los jugadores con el código moral adoptado conjuntamente [también] les impide violar los derechos de los demás jugadores o sus propias obligaciones con el fin de *incrementar la utilidad social*» temporalmente, a expensas de la utilidad social a largo plazo (ibid., p. 694, cursivas originales). Las desviaciones utilitaristas de actos de un código óptimo no son más permisibles que las desviaciones egoístas. Alejan por igual de los beneficios de la libertad, productividad y seguridad que el código óptimo hace posible.¹³

III. PRIORIDADES ABSOLUTAS DENTRO DE UN CÓDIGO UTILITARISTA LIBERAL

El utilitarismo liberal tal como ha sido interpretado está expuesto a varias objeciones, a pesar de su considerable fuerza. In-

¹³ Pueden obtenerse juegos más generales de utilitarismo de reglas relajando las suposiciones de información perfecta y completa y de cumplimiento pleno. Véase, v.gr., Harsanyi (1985b), pp. 72-3; (1992), p. 694.

cluso si se acepta que la utilidad representa meramente la conducta de elección racional, los pensadores morales podrían no aceptar todos los criterios de racionalidad que intervienen en el modelo moral complejo de conducta de elección racional que, para Harsanyi, cabe representar por medio de una función de utilidad moral del tipo utilitarista. Los mismos utilitaristas podrían rechazar, por ejemplo, su condición dworkiniana de igual consideración y respeto para los intereses de cada persona, la cual prohíbe a los observadores impersonales tener en cuenta las preferencias de cualquiera «orientadas a otros», incluidas elecciones benévolas y malévolas (ibid., pp. 704-6). Sin mayor argumentación, la condición parece dar origen a un círculo vicioso: veda al cálculo moral computar siquiera utilidades orientadas a otros, y lo hace por razones morales que no son productos del cálculo sino que vinculan *previamente* a su operación. Hare (1981, pp. 169-96), por lo pronto, en apariencia ve arbitraria esa restricción deontológica sobre la moral utilitarista.

Los no utilitaristas que aceptan la teoría ortodoxa de la utilidad pueden rechazar completamente su modelo de equiprobabilidad de la elección impersonal junto con el postulado de similitud fundamental.¹⁴ Binmore sostiene, por ejemplo, que el modelo en la práctica nos pide «imaginar alguna clase de consenso kantiano [sobre la comparación interpersonal] que existiría *con anterioridad a la sociedad misma*. No sólo eso, también debemos creer que la naturaleza de este consenso kantiano puede ser deducida de alguna forma por un ser suficientemente racional a partir de datos que no hacen referencia al depósito de conocimiento común que [puede identificarse] con la cultura de una sociedad» (1994, p. 299, cursivas originales). Binmore ofrece un modelo alternativo de elección moral, en el cual un consenso sobre comparaciones interpersonales no es el producto de la racionalidad impersonal, sino de la evolución social. La evolución social tiende a favorecer la supervivencia de aquellas comparaciones interpersonales hechas por personas cuyos patrones de conducta se consideren con más éxito que los de otras personas, y así con más probabilidad de ser imitados y alentados en la educación, en un contexto cultural concreto. En un «equilibrio empático» (en el cual nadie tiene un in-

¹⁴ Como es bien sabido, Rawls rechaza completamente la teoría de la utilidad. Él se propone basar sus principios políticos de justicia en «bienes primarios» supuestamente valorados por todas las personas racionales y razonables, a pesar de la pluralidad e inconmensurabilidad de sus sistemas comprensivos de valores.

centivo para apartarse de las comparaciones interpersonales que escoge dadas las comparaciones interpersonales hechas por otras personas), hay un consenso cultural sobre cómo hacer las comparaciones interpersonales de utilidad, sobre la base del cual gente con poder de negociación acaso señaladamente desigual coordina sus interacciones sociales para el momento (1997, capítulo 4).

Sin embargo pasaré por alto en gran medida tales objeciones, para responder a la aversión de Harsanyi hacia las prioridades absolutas dentro de un código utilitarista liberal. Como ya se expuso, el utilitarismo de reglas acoge algo de la prioridad léxica que Rawls da a los derechos básicos iguales, en el sentido de que un observador impersonal debe determinar qué es *practicable* en términos del equilibrio de Nash antes de elegir un código *óptimo*. Se da prioridad absoluta a un conjunto de estrategias permisibles de un código óptimo sobre todas las demás consideraciones. No obstante, ciertamente no hay garantía de que un sistema utilitarista de iguales derechos y deberes deba semejarse mucho al sistema específico defendido por el liberalismo político. Todo lo que sabemos es que los derechos y libertades iguales han de ser distribuidos en cualquier situación en la que se maximice la utilidad social.¹⁵ *A fortiori*, no hay garantía de que un subconjunto apropiado de derechos básicos siempre superará otros derechos y deberes en casos de conflicto. Por el contrario —como insiste Harsanyi—, las reglas de prioridad absoluta dentro de un código utilitarista aparecen irrazonablemente rígidas, ya que la maximización de la utilidad social puede exigir que cualquier derecho pueda dar paso a otros en algunas situaciones.

Con todo, parece posible modificar la teoría de Harsanyi, de forma que surja un utilitarismo de reglas que incorpora prioridades liberales. A este respecto, su rechazo de las reglas de prioridad léxica dentro de un código óptimo está ligada a su concepción del tipo moral de elección racional. Esa concepción implica lo que llamaré «monismo fuerte», la idea de que todos los conflictos de va-

¹⁵ Esa prescripción abstracta, sin embargo, no distingue al utilitarismo de reglas de teorías morales alternativas como la teoría de la negociación proporcional propuesta por Binmore (1994, 1997). A diferencia del utilitarismo, la negociación proporcional puede requerir que a una persona pobre con un nivel de utilidad inicial relativamente bajo se le otorgue el derecho a un beneficio médico tal como un trasplante de órganos, incluso aunque una persona rica con un nivel de utilidad inicial relativamente alto hubiese obtenido del trasplante mayor satisfacción de preferencias comparable. Además, como observa Binmore (1994, pp. 146-7), su teoría no garantiza que los derechos iguales sean distribuidos a todos.

lores pueden en último término resolverse en términos de una sola función de utilidad social homogénea $w = w(\bar{s}, M)$, que refleja una preferencia moral completa y transitiva por los equilibrios de Nash. El monismo fuerte supone la llamativa hipótesis de que los diversos actos escogidos por una persona son todos de la misma clase a efectos normativos, reductibles en último término a una sola escala de valor homogéneo. Pero, más allá de la definición de elección racional y moral que la manifiesta, no se proporciona ninguna justificación para una hipótesis tan singular. De hecho, la ruta tradicional hacia ella (esto es, el crudo hedonismo) es desestimado por el mismo Harsanyi como increíble.

Para garantizar que las prioridades liberales absolutas se incorporan dentro de un código óptimo, los utilitaristas de reglas deben adoptar una teoría de la elección moral que no implica un monismo fuerte. Tal teoría supondrá modificaciones en la serie de condiciones de racionalidad que intervienen en la concepción de Harsanyi de un tipo moral de elección racional. Considérese, por ejemplo, los criterios de utilidad esperada junto con su afirmación característica de que un maximizador de la utilidad esperada debe usar sus actitudes hacia el riesgo (tal como se revelan en forma de una función de utilidad vNM) para medir sus intensidades de preferencia relativas. Según la teoría vNM, una persona racional se comporta en condiciones de riesgo *como si* tuviera no sólo una ordenación de preferencias respecto a sus posibles acciones alternativas, sino también creencias consistentes sobre los sucesos alternativos, donde sus creencias corresponden a una distribución de probabilidad objetiva (idéntica para todas las personas racionales). Las acciones $f \in F$ son vistas como funciones desde estados de cosas $e \in E$ a consecuencias $x \in X$. Una consecuencia concreta x es el efecto conjunto de una acción f y un estado e que sucede con probabilidad $p_i \in p$. Las preferencias sobre consecuencias son un caso especial de preferencias sobre acciones porque, como señala Mongin, «una consecuencia x puede identificarse con el acto que tiene valor constante x » (1997, p. 9, notación alterada). Los axiomas vNM implican que las preferencias de la persona pueden representarse por una función de utilidad esperada $\sum u(f(e))p(e)$, donde $u(\cdot)$ es una función de X , y $p(\cdot)$ es una distribución de probabilidad de E . Esa función de utilidad vNM u es una función cardinal, ya que es única hasta una transformación positiva afín.¹⁶

¹⁶ Para más detalles sobre la teoría vNM, véase Binmore (1994), pp. 266-78, 304-15. Para el caso presente puede dejarse al margen la extensión de la te-

Es sin duda cierto, como han insistido Weymark (1991) y otros, que esta ordenación racional de las preferencias de una persona puede ser representada por igual mediante muchas funciones de utilidad además de las funciones de utilidad vNM. No hay nada en su conducta electora en condiciones de riesgo que justifique una limitación a la familia cardinal de las funciones de utilidad vNM. Pero Harsanyi (1992, p. 685) afirma que «en ética», la racionalidad exige también que la persona emplee sus *actitudes* mentales hacia el riesgo, tal como se incorporan en la forma de una función de utilidad VNM, para medir sus intensidades relativas de preferencias sobre acciones alternativas. Esa condición de racionalidad, en cuanto que difiere de la exigencia de *comportarse* con arreglo a los criterios de utilidad esperada, restringe la representación de tal conducta a las funciones de utilidad vNM. Harsanyi aparentemente ve la condición como un imperativo hipotético de la razón, cuya validez será autoevidente a quienquiera que tome conciencia de las actitudes concretas hacia el riesgo que están latentes en él y le ayudan a definirse como una persona. Pero no hay ningún intento de explicar la *formación* de tales actitudes. Por el contrario, simplemente se asume su posesión por cualquier ser humano racional y moral. Por tanto parecería que Harsanyi sencillamente las saca de la nada, y se basa en ellas sin justificación para introducir de contrabando la cardinalidad que su utilitarismo de reglas necesita para funcionar.¹⁷

Con todo, la afirmación de que una persona moral y racional posee, y hace uso adecuado de, actitudes hacia el riesgo así como otras actitudes [es un] requisito de la ética utilitarista, puede se-

oría que hacen Savage y otros al supuesto de incertidumbre, donde las probabilidades objetivas no están disponibles. Esto no significa negar que es bastante atrevido suponer que todas las personas racionales (incluidos los observadores impersonales) se comportan en condiciones de incertidumbre como si se enfrentasen a una distribución de probabilidad idéntica. Sin embargo, como señala Binmore, la teoría ortodoxa de la utilidad subjetiva considera tal conducta como virtualmente tautológica.

¹⁷ Frente a esto, Binmore sostiene que la *evolución social* selecciona a favor de la conducta consistente que sólo puede representarse como maximización de una función de utilidad VNM. Su idea parece ser que los individuos que tienen éxito tienden a comportarse como si poseyesen actitudes hacia el riesgo que usan para medir intensidades de preferencias relativas. La necesidad de sobrevivir en un proceso competitivo de evolución social explica la restricción exigida a las funciones cardinales de utilidad: «[L]a evolución no premia la conducta consistente porque sea consistente. Sólo premia la conducta consistente que se dirige a su propia supervivencia» (1994, pp. 324-5; véase también *ibid.*, pp. 273-4, n. 21). Evidentemente, la negociación de Nash (que es primordial para el enfoque de Binmore) descansa en esa información de utilidad cardinal.

pararse de la pregunta de cómo una persona llega a poseer las actitudes relevantes. Seguiré a Harsanyi en este punto, con la advertencia de que hace falta decir más cosas acerca de cómo se forman las actitudes relevantes. Algo más se dirá, aunque brevemente, en la conclusión.

Para ir más allá del monismo fuerte tal como ha de hacerse para dar prioridad absoluta a los derechos iguales dentro de un código óptimo, la teoría ortodoxa de la utilidad esperada debe ser reemplazada por lo que podría llamarse una teoría «plural» de la utilidad esperada (cf. Sen, 1980). Según ésta, la conducta racional en condiciones de riesgo implica reconocer que el conjunto practicable de acciones debe ser fraccionado en una pluralidad de subconjuntos o clases, donde las diferentes clases se distribuyen en una jerarquía tal que cualquier clase es inherente o infinitamente más (menos) valiosa que otras bajo (sobre) ella en la jerarquía. Como hay una pluralidad de clases de acciones, hay una pluralidad de sistemas de criterios de utilidad esperada, uno por cada clase de acciones (implícitamente, consecuencias). Defínase $F(\alpha)$ como el conjunto de acciones del tipo α , donde cualquier acción $f(\alpha)$ es una función desde los estados de cosas $e(\alpha) \in E(\alpha)$ a las consecuencias $x(\alpha) \in X(\alpha)$ del tipo α . Similarmente, defínase $F(\beta)$ como el conjunto de acciones del tipo β , y así sucesivamente. Sean $x_1(\alpha)$ y $x_m(\alpha)$ las mejores y peores consecuencias o resultados posibles del tipo α para la persona k , $x_1(\beta)$ y $x_m(\beta)$ las mejores y peores consecuencias o resultados posibles del tipo β para él, y así sucesivamente. Defínase $p^\alpha(z^\alpha_i)$ como un sorteo restringido que ofrece $x_1(\alpha)$ con probabilidad $p^\alpha_1 = z^\alpha_i$, y $x_m(\alpha)$ con probabilidad $p^\alpha_m = 1 - z^\alpha_i$, $i = 1, \dots, m$. Similarmente, defínase $p^\beta(z^\beta_i)$ como un sorteo restringido que ofrece $x_1(\beta)$ con probabilidad $p^\beta_1 = z^\beta_i$, y $x_m(\beta)$ con probabilidad $p^\beta_m = 1 - z^\beta_i$, y así sucesivamente. Entonces, siguiendo a Binmore (1994, pp. 272-3), pueden definirse los axiomas de continuidad, independencia, equivalencia probabilística y optimización para cada clase de acciones.

El axioma de continuidad para las acciones del tipo α , por ejemplo, puede ser enunciado como:

Continuidad- α : Para cada consecuencia $x_i(\alpha) \in X(\alpha)$, existe algún sorteo restringido $p^\alpha(z^\alpha_i) = (x_1(\alpha), z^\alpha_i; x_m(\alpha), 1 - z^\alpha_i)$, $0 \leq z^\alpha_i \leq 1$, tal que la persona k es indiferente entre $x_i(\alpha)$ y $p^\alpha(z^\alpha_i)$.

Este axioma dice que es siempre posible para k hallar alguna probabilidad z^α_i de obtener su mejor resultado posible $x_1(\alpha)$ del tipo α que le llevaría a mostrar indiferencia de la clase superior α entre

cualquier resultado seguro dado $x_i(\alpha)$ y un sorteo restringido $p^\alpha(z_i^\alpha) = (x_1(\alpha), z_i^\alpha; x_m(\alpha), 1-z_i^\alpha)$. Un axioma análogo de continuidad- β puede enunciarse para las acciones del tipo β , y así sucesivamente.

Un axioma de independencia- α puede enunciarse como:

Independencia- α : Suponiendo que una persona k es indiferente entre cualquier $x_i(\alpha)$ y un sorteo restringido $p^\alpha(z_i^\alpha)$, $i = 1, \dots, m$, entonces k es indiferente si los m resultados seguros $x_i(\alpha)$ en cualquier sorteo p^α son reemplazados por los m sorteos restringidos $p^\alpha(z_i^\alpha)$.

En otras palabras, a la persona k no debe importarle si un sorteo simple se transforma en un sorteo compuesto con lo que, para él, son resultados o premios equivalentes del tipo α . De nuevo, puede definirse una axioma de independencia- β análogo, y así sucesivamente.

Sin entrar en mayores detalles, basta subrayar que pueden imponerse distintos axiomas de continuidad, independencia, equivalencia probabilística y optimización en el contexto de cada clase de acciones. Cualquier persona que se comporte de acuerdo con esos criterios plurales de utilidad esperada puede ser descrita como que maximiza las distintas funciones de utilidad VNM u^α, u^β, \dots definidas a partir de las diferentes clases de resultados (seguros), respectivamente. Al mismo tiempo, un maximizador de la utilidad esperada pluralista ha de otorgar prioridad absoluta a las clases superiores de acciones sobre las clases inferiores. Si una acción de una clase superior entra en conflicto con una de otra clase inferior, debe siempre escoger la clase superior. Dejando aparte la cuestión de qué distingue las naturalezas de las diferentes clases de actos, un agente racional de este tipo se comporta como si tuviese una jerarquía correspondiente de distintas ordenaciones de preferencia, una por cada una de la pluralidad de clases de acciones. Además se supone que posee actitudes complejas hacia el riesgo, las cuales emplea para medir las intensidades de sus preferencias relativas de cada clase, y para disponer la pluralidad de clases en un orden léxico (idéntico para todos los agentes racionales). La ordenación léxica refleja la reluctancia de todo agente racional hacia arriesgarse en algún momento a perder cualquier cantidad de una clase de utilidad superior para ganar cualquier cantidad de una clase inferior. Su preferencia por la clase superior de acciones es *infinitamente* más intensa que su preferencia por la clase inferior.

Este tipo de conducta racional puede describirse como la maximización en una jerarquía de clases plurales distintas de funciones de utilidad VNM, una por cada clase de acciones. Cuando escoge entre actos de la misma clase, el sujeto se comporta como si maximizase el valor de una función de utilidad VNM restringida al conjunto de acciones relevantes. Las intensidades de sus preferencias relativas se miden por sus actitudes hacia el resto en el contexto de las acciones de esa clase. Sin embargo, cuando escoge entre acciones de clases diferentes, invariablemente selecciona la clase superior de acción. Se comporta siempre como si emergiese hacia lo alto de la jerarquía cuando puede hacerlo, desiteresándose de sus clases inferiores de funciones de utilidad cuando es posible la maximización de una clase superior.

Si la teoría pluralista de la utilidad esperada se combina adecuadamente con las otras condiciones de la elección moral de Harsanyi (incluidos el criterio de equiprobabilidad y demás), un observador impersonal se comporta como si maximizase respecto de un conjunto de funciones de utilidad social y moral heterogéneas w^α, w^β, \dots , que se relacionan unas con otras por reglas de prioridad léxica, en lugar de respecto de una función homogénea singular $w = w(\bar{s}, M)$. La clase superior de función de utilidad social y moral, $w^\alpha = w^\alpha(\bar{s}^\alpha, M^\alpha)$, representa su juicio acerca de las reglas morales en el contexto de la clase de acciones con mayor valor inherente, las cuales se supone que él (y cualquier otro observador como él) ven como derechos protegidos moralmente. En efecto, el observador maximiza w^α construyendo el mejor sistema coherente de derechos iguales para todos.

Surge así un tipo más complejo de utilitarismo liberal milliano, el cual reconoce la prioridad absoluta de un sistema de derechos básicos iguales sobre todas las demás estrategias permitidas (y no permitidas) por un código óptimo. Por ejemplo, supongamos que hay dos clases diferentes de acciones. Las acciones de la clase superior son tan inherentemente valiosas que un observador impersonal asocia invariablemente a su realización derechos-pretensión morales, donde la pretensión del titular del derecho tiene su correlato en obligaciones morales de otros (ya sean de no interferir con su acción o de actuar en su favor). Las acciones de la clase inferior son vistas por el observador como permisibles, pero insuficientemente valiosas para estar protegidas moralmente. Todos tienen una libertad o permiso para realizar tales acciones en competencia con sus compañeros, y ninguno de ellos tiene una obligación de no interferir con esta clase de acciones de otras personas. Entonces, un código moral M^* consiste en dos ca-

pas de sub-códigos $M^{\alpha*}$, $M^{\beta*}$, relacionados por una regla de prioridad léxica. Similarmente, el conjunto de estrategias permisibles $P(M^*)$ se asemeja a una tarta en capas, pues implica dos capas distintas de estrategias permisibles dispuestas en una jerarquía. En casos de conflicto se otorga prioridad absoluta a la capa superior de derechos-pretensión s^{α} sobre la segunda capa de meras libertades s^{β} . La regla de prioridad absoluta se pone de manifiesto en el hecho de que las obligaciones morales se correlacionan sólo con derechos-pretensión. Todos tienen un deber de no interferir con cualquier elección estratégica personal catalogada por el código como un derecho-pretensión, mientras que no hay deber de no competir con elecciones personales catalogadas como meras libertades. A menos de que se suspenda el código, los derechos-pretensión *sólo* pueden ser postergados por otros derechos-pretensión, nunca por meras libertades.

Permanece abierta la cuestión de si este utilitarismo liberal es más atractivo que la forma más flexible de utilitarismo liberal de Harsanyi. Por ejemplo, hace falta clarificar qué acciones son infinitamente más valiosas que otras, y por qué. También se necesitan modelos más adecuados de derechos-pretensión y meras libertades, para precisar el sentido en el que una clase de estrategias permisibles es más poderosa que la otra. En este sentido, los derechos-pretensión pueden verse como estrategias permisibles que, al exigir a otras personas realizar acciones o inacciones (deberes) correlativos, consolidan o mantienen aspectos personales de resultados en equilibrio de Nash de los juegos practicados en la segunda fase del juego utilitarista de reglas en dos etapas. Tales aspectos personales son aquí puestos en manos del titular del derecho, incluso aunque él no sea siempre el agente de las acciones relevantes (puede que los sujetos a deberes tengan que actuar en su favor). Las meras libertades o permisos, por otro lado, son estrategias permisibles que no imponen deberes correlativos sobre otros. Esta clase de estrategia más débil consolida aspectos personales de los equilibrios de Nash sólo con alguna probabilidad positiva menor que uno, dependiendo del poder de negociación relativo del titular de la libertad. Por ejemplo, dos o más jugadores pueden tener cada uno de ellos libertades para competir por el mismo trabajo escaso, el cual sólo puede conseguir un jugador. Pero hace falta avanzar más en esta área.¹⁸

¹⁸ Ha habido en la literatura una tendencia a ver una marcada dicotomía entre las formulaciones de derechos y libertades con forma de juegos y las de

IV. CONCLUSIÓN

Mi argumentación ha sido doble. En primer lugar, el utilitarismo de reglas liberal parece proporcionar una «base segura y aceptable» para la libertad igual, contrariamente a lo que sugieren Rawls y otros.¹⁹ La variedad de Harsanyi concede prioridad absoluta a algún sistema de derechos y libertades, en virtud del hecho de que un código moral óptimo (y su conjunto permisible de estrategias) tiene prioridad absoluta sobre todas las consideraciones no morales (sección II).

En segundo lugar, el utilitarismo liberal milliano puede incluso incorporar prioridades liberales absolutas *dentro* de su código óptimo, aunque para hacerlo así es necesario modificar la teoría de la utilidad ortodoxa de Harsanyi (sección III).

Para terminar, un reparo legítimo contra los utilitarismos liberales tal como suelen concebirse, ya sea en la variedad de Harsanyi o en la más compleja variedad milliana, es que presuponen algo así como observadores impersonales kantianos, con notables capacidades intelectuales y de empatía para hacer juicios morales y racionales que parecen trascender cualquier contexto cultural concreto. En efecto, un observador impersonal debe haber desarrollado lo que Aristóteles y otros antiguos llamaron «phrónesis», esto es, una habilidad tal en el arte de vivir que su conducta electora exhibe tanto el conocimiento como la motivación que hacen falta para alcanzar una armonía ideal de valores sociales y morales, la cual se representa luego como la maximización de la utili-

elección social. Para el vivo debate vigente, véase v.gr. Gaertner, Pattanaik y Suzumura (1992), Sen (1992), Pattanaik y Suzumura (1994), y Hammond (1995). Sin embargo, como subraya Sen, «la supuesta dicotomía es más de presentación que de sustancia» (1992, p. 153). Los modelos con forma de juego se concentran en las estrategias permisibles de cada persona, mientras que los modelos de elección social se concentran directamente en el poder para influir en los resultados sociales. Pero la definición de permisibilidad es «obtenida —directa o indirectamente— a la luz de las características o consecuencias de combinar las estrategias de diferentes personas» (ibid.). Y una «combinación de estrategias» en equilibrio «produce un resultado —un estado social (aunque se describa el estado social sólo como una combinación concreta de acciones ocurridas)» (ibid.). Para más detalles, véase Riley (1997f).

¹⁹ Riley (1997c) argumenta que el liberalismo político de Rawls, debido a su peculiar estructura, es en sí mismo incapaz de garantizar que los valores asociados a los derechos iguales deban *superar* otros valores opuestos, incluidos valores liberales.

dad social y moral.²⁰ Pero no hay ninguna pista sobre el proceso por el cual se adquiere la *phrónesis*.

Una posible vía de respuesta podría ser elaborar una explicación evolucionista apropiada, como hace Binmore para apoyar su teoría de la negociación proporcional en relación con la elección moral. Tal explicación reposa sobre la noción de aptitud o éxito en un proceso competitivo de evolución social en cuanto motivación del sujeto racional y moral, sin referirse necesariamente a su deseo-satisfacción o a sus sentimientos de placer y dolor.

Otra posible vía sería proporcionar una teoría psicológica para explicar cómo un sujeto podría educarse a sí mismo para llegar a ser un observador impersonal en un amplio ámbito de contextos sociales. Una variante especialmente controvertida de esta vía de «psicología popular» es la del utilitarismo clásico, a saber: el hedonismo. En efecto, a diferencia de Harsanyi, los utilitaristas tanto clásicos como modernos trasladan normalmente su teoría de la motivación directamente al concepto mismo de utilidad, de modo que la utilidad denota placer o deseo-satisfacción. Como consecuencia, los términos de «utilidad» y «preferencia» están rodeados de una gran confusión en la literatura. Por ejemplo, algunos equiparan utilidad y preferencia con deseo-satisfacción, mientras que otros ven la utilidad como una mera representación de cualquier preferencia revelada por una conducta electora racional, con independencia del deseo-satisfacción. Sin embargo, si las jerarquías de preferencias basadas en deseos entran en conflicto con las basadas en elecciones racionales, no es posible que la utilidad se refiera a ambas. Sólo cuando las dos ordenaciones coinciden pueden reconciliarse los dos usos de utilidad y preferencia. En ese caso, el término «utilidad» podría usarse para referirse, simultáneamente, a la motivación última que inspira una jerarquía de preferencias, y al valor de una función que representa esa misma preferencia tal como la revela la conducta racional y moral. En otra parte he defendido un utilitarismo liberal ideal de Mill con esta forma, en el cual los derechos a la libertad de acción «puramente referida a uno mismo» son vistos como pretensiones absolutas, que no pueden ser desplazadas por ninguna consideración

²⁰ Para una aproximación a la *phrónesis*, véase Annas (1993), pp. 73-84. Como ella explica, «la *phrónesis* unifica las virtudes» y «la persona plenamente virtuosa, con posesión completa de *phrónesis*, es un ideal» (ibid., p. 83). Por desgracia, ella parece pensar que «el razonamiento maximizador» es incompatible con la *phrónesis*, y más en general con la teoría moral antigua (ibid.), pp. 86-7, 447-8.

opuesta (incluidos otros derechos), con la importante advertencia de que debe definirse la clase de acciones puramente referidas a uno mismo.²¹

Jonathan Riley
Murphy Institute of Political Economy
Tulane University
New Orleans, La. 70118-5624 (USA)
email: jonriley@mailhost.tcs.tulane.edu

REFERENCIAS

- ANNAS, Julia (1993). *The Morality of Happiness*. New York and Oxford: Oxford University Press.
- BINMORE, Ken (1994). *Playing Fair*. Cambridge: The MIT Press.
- , (1997). *Just Playing*. Cambridge: The MIT Press. Los números de página se refieren al borrador manuscrito.
- BROOME, John (1991). *Weighing Goods*. Oxford: Blackwell.
- DWORKIN, Ronald (1977). *Taking Rights Seriously*. London: Duckworth. [Hay versión española de M. Gustavino: *Los derechos en serio*, Barcelona: Ariel, 1984].
- GAERTNER, Wulf, PATTANAIK, Prasanta, y SUZUMURA, Kotaro (1992). «Individual rights revisited», *Economica* 59: 161-77.
- HAMMOND, Peter J. (1995). «Game forms versus social choice rules as models of rights», artículo en elaboración.
- HARE, Richard M. (1981). *Moral Thinking*. Oxford: Clarendon Press.
- HARSANYI, John C. (1977). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.
- , (1982). «Morality and the theory of rational behavior», in A.K. Sen and B. Williams, eds., *Utilitarianism and Beyond*. Cambridge: Cambridge University Press. Pp. 39-62. Publicado originalmente en *Social Research* 44 (1977): 623-56.
- , (1985a). «Does Reason Tell Us What Moral Code to Follow and, Indeed, to Follow Any Moral Code At All?», *Ethics* 96: 42-55.

²¹ Véase Riley (1988), (1989-90), (1997a), (1997b), (1997g).

- , (1985b). «On Preferences, Promises and the Coordination Problem», *Ethics* 96: 68-73.
- , (1992). «Game and decision theoretic models in ethics», en R.J. Aumann and S. Hart, eds., *Handbook of Game Theory*, I. Amsterdam: North-Holland. Pp. 669-707.
- , (1996). «John Rawls's Theory of Justice: Some Critical Comments», en este volumen. Los números de página se refieren al borrador de agosto.
- MILL, John Stuart (1859). *On Liberty*. London: Parker. [Hay versión española de P. de Azcárate: *Sobre la libertad*, Madrid: Alianza, 1986].
- , (1861). «Utilitarianism», *Fraser's Magazine* 64: 391-406, 525-34, 658-73. [Hay versión española de E. Guisán: *El utilitarismo*, Madrid: Alianza, 1984].
- MONGIN, Philippe (1997). «Spurious Unanimity and the Pareto Principle», artículo en elaboración. Page references are to February draft.
- PATTANAIK, Prasanta, y SUZUMURA, Kotaro (1994). «Rights, welfare and social choice», *American Economic Review: Papers and Proceedings* 84: 435-9.
- RAWLS, John (1971). *A Theory of Justice*. Cambridge: Harvard University Press. [Hay versión española de M^a D. González: *Teoría de la justicia*, Madrid: Fondo de Cultura Económica, 1979].
- , (1985) «Justice as Fairness: Political, not Metaphysical», *Philosophy & Public Affairs* 13: 225-251.
- , (1993). *Political Liberalism*. New York: Columbia University Press. [Hay versión española de A. Domènech: *El liberalismo político*, Barcelona: Crítica, 1996].
- , (1995). «Reply to Habermas», *Journal of Philosophy* 92: 132-80.
- RILEY, Jonathan (1988). *Liberal Utilitarianism*. Cambridge: Cambridge University Press.
- , (1989-90). «Rights to liberty in purely private matters, parts I-II», *Economics and Philosophy* 5 & 6: 121-66, 27-64.
- , (1996). «Mill on justice», in D. Boucher and P.J. Kelly, eds., *Social Justice: From Hume to Hayek*. London: Routledge. De próxima publicación.
- , (1997a). *Philosophical GuideBook to J.S. Mill: On Liberty*. London: Routledge. De próxima publicación.
- , (1997b). *Absolute Liberty of Purely Self-Regarding Action: A Defense of Mill's Doctrine*. London: Routledge. De próxima publicación.
- , (1997c). «A Difficulty in Rawl's Conception of Political Liberalism», artículo en elaboración.
- , (1997d). «Interpreting Harsanyi's Rational Choice Utilitarianism», artículo en elaboración.

- , (1997e). «Should Rule Utilitarians Ignore Other-Oriented Preferences?», artículo en elaboración.
- , (1997f). «Modelling Rights and Liberties», artículo en elaboración.
- , (1997g). *Security of Equal Rights: A Liberal Utilitarian Theory of Justice*. Manuscrito borrador.
- SEN, Amartya K. (1980). «Plural utility», *Proceedings of the Aristotelian Society* 81: 193-215.
- , (1992). «Minimal liberty», *Economica* 59: 139-59.
- WEYMARK, John A. (1991). «A reconsideration of the Harsanyi-Sen debate on utilitarianism», en J. Elster y J. Roemer, eds., *Interpersonal Comparisons of Well-Being*. Cambridge: Cambridge University Press. Pp. 255-320.