VI. LOS TONTOS RACIONALES: UNA CRÍTICA DE LOS FUNDAMENTOS CONDUCTISTAS DE LA TEORÍA ECONÓMICA*

AMARTYA K. SEN

I

EN SU *Mathematical Psychics,* publicado en 1881, afirmaba Edgeworth que "el primer principio de la Economía es que cada agente está movido sólo por su propio interés".[1] Esta concepción del hombre ha sido persistente en los modelos económicos, y la naturaleza de la teoría económica parece haberse visto muy influida por esta premisa básica. En este ensayo me gustaría examinar algunos de los problemas surgidos de esta concepción de los seres humanos.

Tomado de *Philosophy and Public Affairs,* 6, 1976-1977, pp. 317-344.

* Esta conferencia de Herbert Spencer, pronunciada en la Universidad de Oxford en octubre de 1976, aparecerá en *Scientific Models and Man,* comp. H. Harris (próxima publicación, 1978) y se reproduce aquí gracias al amable permiso concedido por la Oxford University Press. Agradezco los útiles comentarios formulados sobre una versión anterior por los directores de esta revista y por Åke Anderson, Isaiah Berlin, Frank Hahn, Martin Hollis, Janos Kornai, Derek Parfit, Christopher Peacocke y Tibor Scitovsky.

[1] F. Y. Edgeworth, *Mathematical Psychics. An Essay on the Application of Mathematics to the Moral Sciences* (Londres, 1881), p. 16.

Mencionaré que el propio Edgeworth estaba bien consciente de que su primer principio de la economía no era particularmente realista. En efecto, creía que "el hombre concreto del siglo XIX es en su mayor parte un egoísta impuro, un utilitarista mixto".[2] Esto plantea el interesante enigma de la razón de que Edgeworth dedicara tanto tiempo y talento a desarrollar una línea de investigación cuyo primer principio consideraba falso. No se trata de saber por qué deban emplearse abstracciones en la investigación de las cuestiones económicas generales —la naturaleza de la investigación hace que esto sea inevitable—, sino por qué habremos de escoger un supuesto que no sólo consideramos incorrecto en detalle sino fundamentalmente equivocado. Como podremos observar esta cuestión tiene un interés continuo también para la economía moderna.

Por lo que toca a Edgeworth, una parte de la respuesta residía sin duda alguna en el hecho de que no consideraba fundamentalmente errado el supuesto en los tipos de actividades *particulares* a los que aplicaba lo que llamó el "cálculo económico": i) la guerra y ii) el contrato. "Admitiendo que existe en las partes superiores de la naturaleza humana una tendencia hacia las instituciones utilitarias y un aprecio por tales instituciones", se planteó este interrogante retórico: "podríamos su-

[2] Edgeworth, p. 104. En efecto, más adelante hizo algunas observaciones interesantes sobre los resultados del egoísmo "impuro", admitiendo un elemento de simpatía hacia los demás. Las observaciones han sido investigadas y analizadas por David Gollard en "Edgeworth's Propositions on Altruism", *Economic Journal,* 85 (1975).

poner seriamente que estas consideraciones morales son aplicables a la guerra y el comercio, que podrían erradicar el núcleo incontrolado del egoísmo humano, o ejercer una fuerza apreciable por comparación con el impulso del interés propio".[3] En su opinión, Sidgwick había despejado la "ilusión" de que "el interés de todos es el interés de cada uno", ya que descubrió que "los dos principios supremos —Egoísmo y Utilitarismo—" son "irreconciliables, a no ser por la acción de la religión". "Lejos está del espíritu de la filosofía del placer la subestimación de la importancia de la religión", escribió Edgeworth, "pero en la investigación actual, y tratándose de los elementos inferiores de la naturaleza humana, debemos tratar de buscar una transición más obvia, un pasaje más terrenal, a partir del principio del interés propio, para llegar al principio, o por lo menos la práctica, del utilitarismo".[4]

Adviértase que el contexo del debate es importante para este argumento. Edgeworth pensó que había establecido la aceptabilidad del "egoísmo" como el supuesto conductista fundamental para su investigación particular, destruyendo la aceptabilidad del "utilitarismo" como una descripción del comportamiento efectivo. Por supuesto, el utilitarismo no es el único enfoque no egoísta. Además, entre los intereses propios y los intereses de todos se encuentran los intereses de diversos grupos: familias, amigos, comunidades locales, grupos de colegas, clases económicas y sociales. Los conceptos de la responsabilidad familiar, la ética empresarial,

[3] Edgeworth, p. 52.
[4] *Ibid.*, pp. 52-53.

la conciencia de clase, etc., se refieren a estas áreas de interés intermedias, y la eliminación del utilitarismo como una teoría descriptiva del comportamiento no deja al egoísmo como la única opción. Sería difícil negar la importancia de algunas de estas consideraciones para la economía de las negociaciones y los contratos.

Debe advertirse que la investigación de Edgeworth acerca del resultado del contacto económico entre individuos puramente egoístas tuvo el mérito de ser inmediatamente pertinente para una discusión abstracta que ya se prolongaba por más de un siglo y en cuyos debates participaron Herbert Spencer, Henry Sidgwick y otros pensadores destacados del periodo. Dos años antes de la aparición del *Mathematical Psychics* de Edgeworth, había publicado Herbert Spencer su meticuloso análisis de la relación existente entre el egoísmo y el altruismo en *The Data of Ethics*. Había llegado a la conclusión reconfortante —aunque algo confusa— de que "la felicidad general tendrá que alcanzarse principalmente mediante la búsqueda adecuada de su propia felicidad por parte de los individuos; al mismo tiempo que la felicidad de los individuos tendrá que alcanzarse en parte mediante su búsqueda de la felicidad general".[5] En el contexto de esta noción relativamente abstracta, el preciso análisis económico de Edgeworth, basado en un modelo bien definido de los contratos celebrados entre dos individuos egoístas, o entre dos tipos de individuos egoístas (idénticos), ofrecía una respuesta clara para una antigua cuestión hipotética.

[5] H. Spencer, *The Data of Ethics* (Londres, 1879; edición aumentada, 1887), p. 238.

En el modelo de Edgeworth, basado en el comportamiento egoísta, parecía existir una correspondencia notable entre los equilibrios del intercambio de los mercados competitivos y lo que en términos económicos modernos se llama "el núcleo" de la economía. Se dice que un resultado se encuentra en "el núcleo" de la economía si, y sólo si, satisface varias condiciones de imposibilidad de mejoramiento. Estas condiciones establecen, en términos generales, que nadie podría mejorar su situación sin que empeorara la de alguna otra persona (la situación llamada "óptimo de Pareto"), pero también que nadie está en peor situación de la que estaría en ausencia de intercambio, y que ninguna coalición de individuos podría mejorar por sí sola su propia situación alterando el intercambio entre ellos.

Edgeworth demostró que, dados ciertos supuestos generales, cualquier equilibrio que pueda surgir en un mercado competitivo deberá satisfacer estas condiciones y encontrarse en "el núcleo". Así pues, en el modelo de Edgeworth no se encuentran los equilibrios del mercado competitivo dominados en este sentido por ningún otro arreglo viable, dada la distribución inicial de las dotaciones. Más sorprendente era en cierta forma el resultado contrario: que si aumentara el número de individuos de cada tipo ilimitadamente, el núcleo (representativo de tales resultados no superados) se contraería hacia el conjunto de equilibrios competitivos; es decir, el núcleo no sería mucho más extenso que el conjunto de equilibrios competitivos. Este par de resultados se ha refinado y ampliado grandemente en la bibliografía reciente del equilibrio

general con modelos similares y esencialmente con los mismos supuestos conductistas.[6]

Pero el hecho de encontrarse en el núcleo no es por sí solo un gran logro desde el punto de vista del bienestar social. Una persona que empiece mal dotada puede seguir siendo pobre aun después de las transacciones, y si la competencia no ofrece más que la permanencia en el núcleo, la persona desamparada no considerará "la gran cosa" este resultado. Edgeworth tomó en cuenta esta situación hasta cierto punto al considerar el problema de la elección entre diversos equilibrios competitivos. Observó que, para la buena sociedad utilitarista, "la competencia requiere el complemento del arbitraje, y la base del arbitraje entre los contratantes egoístas es la mayor suma posible de la utilidad total".[7] No se ocupó realmente de los aspectos institucionales de tal arbitraje, ni de sus extensas implicaciones para la distribución de la propiedad, a pesar de algunas apariencias superficiales en contrario. Sobre la base del logro de la competencia, por limitado que fuese, Edgeworth se sintió autorizado a "sesgarse hacia una cautela más conservadora en la reforma". Al calcular "la utilidad de las instituciones preutilitaristas", Edgeworth se sintió impresionado "con una visión de la Naturaleza, no totalmente mala como la que nos legara Mill, sino una primera aproximación hacia lo mejor".[8]

En este ensayo no examinaré si la aproximación

[6] Véase en particular K. J. Arrow y F. H. Hahn, *General Competitive Analysis* (San Francisco, 1971). [Versión castellana del FCE, 1977.]

[7] Edgeworth, p. 56.

[8] *Ibid.*, p. 82.

es más bien remota. (Creo que así ocurre aun dentro de la estructura de supuestos utilizada por Edgeworth, pero esto no es fundamental para el tema de este ensayo.) Aquí me interesa la visión del hombre que forma parte del análisis de Edgeworth y sobrevive más o menos intacta en gran parte de la teoría económica moderna. Por supuesto, esta visión es estilizada y se destina específicamente a la iluminación de una disputa relativamente abstracta que interesó mucho a Spencer, Sidgwick y varios otros pensadores contemporáneos prominentes, a saber: ¿en qué sentido y hasta qué punto alcanzaría el bien general el comportamiento egoísta? Desde luego, el hecho de que el comportamiento egoísta sea un supuesto realista no influye en ninguna medida sobre la corrección de la respuesta de Edgeworth al interrogante planteado. Dentro de la estructura de un modelo económico limitado, proveía una respuesta clara a la discusión abstracta del egoísmo y el bien general.

Este debate particular ha persistido durante largo tiempo y aun ahora motiva muchos ejercicios de la teoría económica. La naturaleza limitada del debate ha ejercido una influencia decisiva sobre la elección de modelos económicos y la concepción de los seres humanos que participan en ellos. En su distinguido texto sobre la teoría del equilibrio general, afirman Arrow y Hahn (pp. vi-vii):

> Existe ya una extensa e imponente línea de economistas, desde Adam Smith hasta el presente, que han tratado de demostrar que una economía descentralizada, motivada por el egoísmo y guiada por las señales de los precios, sería compatible con una disposición coherente de recursos económicos que pudiera considerarse, en un sentido bien definido, como superior a la gran clase de disposiciones alternativas posibles. Además, las señales de los precios funcionarían para establecer este grado de coherencia. Es importante entender cuán sorprendente ha de resultar esta afirmación para quien no esté familiarizado con la tradición. La respuesta inmediata "de sentido común" a este interrogante: "¿Cómo sería una economía motivada por la ambición individual y controlada por un número muy grande de agentes diferentes?", será probablemente "el caos". El hecho de que se haya ofrecido durante largo tiempo una respuesta muy diferente, la que en efecto ha impregnado el pensamiento económico de gran número de personas que no son economistas, es razón suficiente para investigarla seriamente. En virtud de que la proposición ha sido planteada y sostenida con gran seriedad, es importante saber no sólo si *es* cierta, sino si *podría* ser cierta. Gran parte de lo que sigue se refiere a este último interrogante, el que merece ampliamente la atención de los economistas en nuestra opinión.

No nos interesa aquí primordialmente la relación entre los modelos postulados y el mundo económico real, sino la corrección de las respuestas a interrogantes bien definidos, planteados con supuestos preseleccionados que constriñen severamente la naturaleza de los modelos que pueden admitirse en el análisis. En el propio interrogante está incrustado un concepto específico del hombre, y no hay libertad para alejarse de esta concepción mientras tratemos de contestar este interrogante. Así pues, la naturaleza del hombre en estos modelos económicos corrientes sigue reflejando la for-

mulación particular de ciertas cuestiones filosóficas generales planteadas en el pasado. El realismo de la concepción escogida del hombre no forma parte de esta investigación.

II

Hay otra razón no empírica —y posiblemente más simple— para que la concepción del hombre en los modelos económicos tienda a ser la de un egoísta que persigue su propio interés. Podemos definir los intereses de una persona en forma tal que parezca perseguir sus propios intereses en cada acto de elección aislado, independientemente de lo que haga.[9] Este enfoque ha sido hasta cierto punto formalizado en época reciente, en el contexto de la teoría de la preferencia revelada, pero tiene una antigüedad respetable, y Joseph Butler ya lo estaba combatiendo en la Capilla Rolls hace dos siglos y medio.[10] La reducción del hombre a la categoría de un animal egoísta depende en este enfoque de una definición cuidadosa. Si se observa que usted escoge x y rechaza y, se declara que usted tiene una preferencia "revelada" por x sobre y. Su utilidad

[9] Si las acciones actuales de una persona afectan su bienestar en el futuro, de acuerdo con este enfoque deberán definirse sus intereses futuros en términos de la forma en que *se evalúan ahora*. En general, no hay razón para presumir que los intereses futuros, tal como se evalúan hoy, coincidirán con tales intereses evaluados en el futuro. Esto aporta una dimensión adicional al problema, y agradezco a Derek Parfit que me haya convencido de la importancia conceptual de esta cuestión.

[10] J. Butler, *Fifteen Sermons Preached at the Rolls Chapel* (Londres, 1726); véase también a T. Nagel, *The Possibility of Altruism* (Oxford, 1970), p. 81.

personal se define entonces simplemente como una representación numérica de esta "preferencia", asignando una utilidad mayor a una opción "preferida". Con este conjunto de definiciones, usted no podrá dejar de maximizar su propia utilidad, excepto por obra de la inconsistencia. Por supuesto, si escoge usted x y rechaza y en una ocasión, y de pronto pasa al polo contrario, el teórico de la preferencia revelada no podrá asignarle un ordenamiento de preferencias, de modo que no podrá asignarle una función de utilidad que usted deba maximizar. Entoncess tendrá que concluir que usted es inconsistente o que sus preferencias están cambiando. También podrá usted frustrar al teórico de la preferencia revelada a través de inconsistencias más refinadas.[11] Pero si usted es consistente, independientemente de que sea usted un egoísta recalcitrante o un altruista delirante o un clasista, aparecerá maximizando su propia utilidad en este mundo encantado de las definiciones. Según la terminología usada en conexión con la tributación, si la justificación del supuesto del egoísmo presentada por Arrow y Hahn equivale a *evitar* el problema, el enfoque de la preferencia revelada parece más un ejemplo sólido de *evasión*.

Este enfoque del egoísmo definicional recibe a veces el nombre de elección racional, e implica nada menos que la consistencia interna. En este enfoque se consideran "racionales" las elecciones de una persona si, y sólo si, *todas* estas eleccio-

[11] Véase H. S. Houthakker, "Revealed Preference and the Utility Function", *Economica*, 17 (1950); P. A. Samuelson, "The Problem of Integrability in Utility Theory", *Economica*, 17 (1950).

nes pueden explicarse en términos de alguna relación de preferencia consistente con la definición de la preferencia revelada, es decir, si todas sus elecciones pueden explicarse como la elección de opciones "preferidas por encima de todas" con respecto a una relación de preferencia postulada.[12] La justificación de este enfoque parece basarse en la idea de que sólo puede entenderse la preferencia real de una persona si se examinan sus elecciones efectivas, y no puede entenderse la actitud de una persona hacia las opciones en forma independiente de la elección. (Por cierto, esta opinión no se limita a los economistas. Hace muchos años, cuando hube de presentar mi examen de literatura inglesa en la Universidad de Calcuta, una de las preguntas que debíamos contestar acerca de *Sueño de una noche de verano* era ésta: Compare los caracteres de Hermia y de Elena. ¿A quién escogería?)

En otra parte he tratado de demostrar que, una vez que hacemos a un lado las curiosas definiciones de la preferencia y el bienestar, este enfoque presupone a la vez demasiadas cosas y demasiado poco: esto último porque hay algunas fuentes de información sobre la preferencia y el bienestar diferentes de las elecciones, tal como suelen entenderse estos términos; y lo primero porque la elección puede reflejar una transacción entre diversas consideraciones, una de las cuales puede ser el bienestar personal.[13]

[12] En cuanto a los principales resultados analíticos, véase M. K. Richter, "Rational Choice", *Preference, Utility and Demand Theory*, comps. J. S. Chipman *et al.* (Nueva York, 1971).

[13] A. K. Sen, "Behaviour and the Concept of Preference", *Economica*, 40 (1973). Véase también el importante estudio reciente de Körner, *Experience and Conduct* (Cambridge, 1971). También T. Schwartz, "Von Wright's Theory of Human Welfare: A Critique", próxima publicación en P. A. Schlipp (comp.), *The Philosophy of Georg Henrik von Wright;* T. Majumdar, "The Concept of Man in Political Economy and Economics", mimeografiado (Universidad Jawaharlal Nehru, Nueva Delhi, 1976); y F. Schick, "Rationality and Sociality", mimeografiado (Universidad Rutgers, Asociación de Filosofía de la Ciencia, 1976).

Varios estudios penetrantes que se ocupan de las decisiones de los consumidores[14] y las actividades productivas[15] han revelado vigorosamente los complejos problemas psicológicos que se encuentran detrás de la elección. No se ha puesto en claro que estas características del comportamiento pue-

[14] Véase T. Scitovsky, *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction* (Londres y Nueva York, 1976). Véase también la crítica general del supuesto del comportamiento "racional" del consumidor, en J. Kornai, *Anti-Equilibrium* (Amsterdam y Londres, 1971), cap. 11; y la bibliografía sobre los "modelos psicológicos de la elección", en particular D. McFadden, "Economic Applications of Psychological Choice Models" (presentado en el Tercer Congreso Econométrico Mundial, agosto de 1975).

[15] Véase H. Liebenstein, "Allocative Efficiency vs. x-Efficiency", *American Economic Review*, 56 (1966). También las críticas formuladas contra el supuesto tradicional de la maximización del beneficio en el comportamiento *empresarial*, en particular W. J. Baumol, *Business Behavior, Value and Growth* (Nueva York, 1959); R. Marris, *The Economic Theory of Managerial Capitalism* (Londres, 1964); O. Williamson, *The Economics of Discretionary Behavior* (Chicago, 1967), y A. Silberston, "Price Behaviour of Firms", *Economic Journal*, 80 (1970), reproducido en Royal Economic Society, *Surveys of Applied Economics*, vol. 1 (Londres, 1973).

dan captarse dentro de los límites formales de la elección consistente de los que depende el enfoque de la maximización del bienestar.[16]

III

Paul Samuelson ha señalado que muchos economistas "separarían la economía de la sociología sobre la base del comportamiento racional o irracional, cuando estos términos se definen en la penumbra de la teoría de la utilidad".[17] Los sociólogos tendrían buenas razones para rechazar esta concepción, pero la cruz que los economistas deben cargar en esta visión de la dicotomía puede advertirse si señalamos que el enfoque del "comportamiento racional", tal como suele interpretarse, conduce a una teoría notablemente muda. Vemos aquí que el comportamiento deberá "explicarse en términos de preferencias, las que a su vez se definen sólo por el comportamiento". Como sería de esperarse, han sido frecuentes los razonamientos circulares. Sin embargo Samuelson tiene razón sin duda cuando afirma que la teoría *"no carece de sentido* en

[16] Acerca de las condiciones de consistencia requeridas para ver la elección en términos de una relación binaria, véase mi ensayo "Choice Functions and Revealed Preference", *Review of Economic Studies,* 38 (1971); H. G. Herzberger, "Ordinal Preference and Rational Choice", *Econometrica,* 41 (1973); K. Suzumura, "Rational Choice and Revealed Preference", *Review of Economic Studies,* 43 (1976); S. Kanger, "Choice Based on Preference", mimeografiado (Universidad de Uppsala, 1976).

[17] P. A. Samuelson, *The Foundation of Economics* (Cambridge, Mass., 1955), p. 90.

términos técnicos".[18] Esto se explica muy simplemente. Como ya hemos señalado, el enfoque impone el requerimiento de la consistencia interna de la elección observada, lo que podría ser refutado por las observaciones efectivas, lo que volvería "significativa" la teoría en el sentido buscado por el enunciado de Samuelson.

El requerimiento de la consistencia tiene una fecundidad sorprendente. Varias características generales de las relaciones de demanda pueden derivarse de él. Pero en el contexto actual, la cuestión principal es la posibilidad de utilizar el requerimiento de consistencia para la *verificación* efectiva. Samuelson especifica la necesidad de "condiciones de observación ideales" para que las implicaciones del enfoque puedan ser "refutadas o verificadas". Sin embargo, esto no se satisface fácilmente porque, por una parte, nuestro amor por la diversidad vuelve ilegítima la consideración de los actos de elección individuales como las unidades apropiadas (en lugar de las *secuencias* de elecciones), mientras que, por otra parte, el paso del tiempo dificulta la distinción entre las inconsistencias y el cambio de los gustos. En efecto, ha habido muy pocos esfuerzos sistemáticos de verificación de la consistencia del comportamiento diario de los individuos, aunque se han realizado algunos experimentos interesantes y útiles sobre las reacciones de los individuos ante la incertidumbre bajo las condiciones del laboratorio. Todavía no se determina lo que haya de considerarse como prueba admisible. Si realizáramos ahora una encuesta entre economistas de diversas escuelas, encontraría-

[18] *Ibid.,* p. 91.

mos casi seguramente una coexistencia de creencias en el sentido de que i) la teoría del comportamiento racional es irrefutable, ii) es refutable y no ha sido refutada hasta ahora, y iii) es refutable y en efecto patentemente falsa.[19]

Sin embargo, ésta no es la cuestión fundamental para los propósitos que persigo aquí. Aun si se obtuviera la consistencia requerida, quedaría insoluta la cuestión del egoísmo, excepto en el sentido puramente definicional, como ya he señalado. Una persona que escoja consistentemente puede tener cualquier grado de egoísmo que deseemos especificar. Por supuesto, es cierto que en el caso especial de la elección pura del consumidor entre los bienes privados, el teórico de la preferencia revelada trata de relacionar la "preferencia" o la "utilidad" de la persona con su *propio* conjunto de bienes. Pero esta restricción no deriva del hecho de que sólo le preocupen sus propios intereses, sino de que su propio conjunto de consumo —o el de su familia— es el único sobre el que ejerce *control* directo en sus actos de elección. La cuestión del egoísmo queda completamente insoluta.

Creo que la cuestión requiere también una formulación más clara que la que suele recibir, y ahora me ocuparé de este punto.

[19] Véanse las críticas filosóficas formuladas recientemente contra la teoría del comportamiento racional en M. Hollis y E. J. Nell, *Rational Economic Man* (Cambridge, 1975); S. Wong, "On the Consistency and Completeness of Paul Samuelson's Programme in the Theory of Consumer Behaviour" (tesis doctoral, Universidad de Cambridge, 1975, próxima publicación). Véanse también las críticas pragmáticas de Kornai en *Anti-Equilibrium*, cap. 11.

IV

Cuando consideramos las desviaciones del "aislamiento poco amable que se supone abstractamente en la Economía", como decía Edgeworth, debemos distinguir entre dos conceptos separados: i) la simpatía y ii) el compromiso. El primero corresponde al caso en el que el interés por otros afecta nuestro propio bienestar. Si el conocimiento de que se tortura a otros nos enferma, éste es un caso de simpatía; si no nos enferma pero creemos que es algo malo y estamos dispuestos a hacer algo para detenerlo, éste es un caso de compromiso. No quiero decir que las palabras escogidas tengan gran mérito, pero me parece que la distinción es importante. Puede sostenerse que el comportamiento basado en la simpatía es egoísta en un sentido importante, porque nos complace el placer de otros y nos duele el dolor de otros, de modo que la búsqueda de nuestra propia utilidad puede ser impulsada por la acción de simpatía. Es la acción basada en el compromiso, no en la simpatía, la que sería altruista en este sentido. (Sin embargo, adviértase que la *existencia* de la simpatía no implica que la acción útil para otros deba *basarse en* la simpatía en el sentido de que no ocurriría si nos complaciera poco o nada el bienestar de otros. Esta cuestión de la *causación* será examinada más adelante.)

En ciertos sentidos, la simpatía es un concepto más fácil de analizar que el del compromiso. Cuando el sentimiento de bienestar de una persona es psicológicamente dependiente del bienestar de otro, se trata de un caso de simpatía; en igualdad de

otras circunstancias, la conciencia del aumento del bienestar de la otra persona mejora entonces directamente la posición de esta persona. (Por supuesto, cuando la influencia es negativa, debemos hablar de una relación de "antipatía", pero podemos ahorrar la terminología y utilizar sólo el término "simpatía", señalando que la relación puede ser positiva o negativa.) Mientras que la simpatía relaciona cosas similares entre sí —el bienestar de diferentes personas—, el compromiso relaciona la elección con los niveles de bienestar esperados. Podemos definir el compromiso en el sentido de que una persona escogerá un acto que en su opinión producirá un nivel de bienestar personal para él menor que otro acto también a su alcance. Adviértase que la comparación se hace entre niveles de bienestar *esperados,* de modo que esta definición del compromiso excluye los actos contrarios al interés propio que derivan sólo de un error en el pronóstico de las consecuencias.

Surge un problema más difícil cuando la elección de una persona coincide con la maximización de su bienestar personal esperado, pero ésa no es la *razón* de su elección. Si queremos incluir este caso, podemos expandir la definición del compromiso para incluir los casos en que la elección de la persona maximiza el bienestar personal esperado pero no cambiará por lo menos bajo una condición contraria en la que el acto escogido cesaría de maximizar el bienestar personal. Es posible que el compromiso en este sentido más amplio resulte difícil de determinar, no sólo en el contexto de las elecciones de otros sino también en el de la propia elección, ya que no siempre está claro lo que hu-

biéramos hecho si las circunstancias hubiesen sido diferentes. Es posible que este sentido más amplio tenga una pertinencia particular cuando actuamos sobre la base de una preocupación por el deber cuya violación podría provocar remordimiento, pero la acción se escoge realmente por el sentido del deber y no sólo para evitar el malestar resultante del remordimiento que ocurriría si actuáramos de otro modo. (Por supuesto, aun el sentido más estrecho del compromiso incluirá el caso en el que el malestar resultante del remordimiento se ve *superado* por el incremento del bienestar.)

Todavía no me he referido a la incertidumbre referente al bienestar esperado. Cuando se introduce la incertidumbre, el concepto de la simpatía no se ve afectado, pero el compromiso requerirá una reformulación. Las modificaciones necesarias dependerán de la reacción de la persona ante la incertidumbre. El caso más simple es probablemente aquel en que la idea que tiene una persona de lo que le ofrece una "lotería" en términos de la ganancia personal se capta en la "utilidad esperada" del bienestar personal (es decir, la adición del bienestar personal derivado de diferentes resultados ponderados por la probabilidad de presentación de cada resultado). En este caso, toda la exposición se reformula simplemente sustituyendo el bienestar personal por el bienestar personal *esperado;* el compromiso implica entonces la elección de una acción que genere un bienestar esperado menor que otra acción disponible. (El sentido más amplio también puede modificarse correspondientemente.)

En la terminología de la teoría económica mo-

derna, la simpatía es un caso de "exterioridad". Muchos modelos descartan las exterioridades; así ocurre, por ejemplo, con el modelo tradicional para el establecimiento de que cada equilibrio competitivo es un óptimo de Pareto y pertenece al núcleo de la economía. Si se permitiera en estos modelos la existencia de la simpatía, se modificarían algunos de estos resultados tradicionales, pero no todos ellos.[20] Pero esto no requeriría una revisión profunda de la estructura básica de estos modelos. Por otra parte, el compromiso implica, en un sentido muy real, la elección en contra de las preferencias, lo que destruye el supuesto crucial de que una opción escogida debe ser mejor que las otras (o por lo menos tan buena como ellas) para que la persona la escoja; y esto requeriría ciertamente que los modelos se formularan de una manera en esencia diferente.

El contraste entre la simpatía y el compromiso puede ilustrarse con la historia de dos niños que se encuentran dos manzanas, una grande y una pequeña. El niño *A* le dice al niño *B*: "Escoge tú". El niño *B* escoge de inmediato la manzana más grande. El niño *A* se enoja y afirma que aquello ha sido muy injusto. "¿Por qué?", pregunta *B*, "¿Cuál habrías escogido *tú*?" "La más pequeña,

[20] Véase A. K. Sen, "Labour Allocation in a Co-operative Enterprise", *Review of Economic Studies*, 33 (1966); S. G. Winter, Jr., "A Simple Remark on the Second Optimality Theorem of Welfare Economics", *Journal of Economic Theory*, 1 (1969); Collard, "Edgeworth's Propositions"; G. C. Archibald y D. Donaldson, "Non-paternalism and Basic Theorems of Welfare Economics", *Canadian Journal of Economics*, 9 (1976).

por supuesto", replica *A*. Entonces *B* concluye con voz triunfante: "¿Por qué te quejas? ¡Ésa es la que te tocó!" El niño *B* gana sin duda esta parte de la discusión, pero en efecto *A* no habría perdido nada por la elección de *B* si su propia elección hipotética de la manzana más pequeña se hubiese basado en la simpatía y no en el compromiso. El enojo de *A* indica que tal vez no hubiera ocurrido así.

Por supuesto, el compromiso se conecta estrechamente a nuestra moral. Pero se trata de la moral en un sentido muy amplio, que abarca diversas influencias, desde la religiosa hasta la política, desde lo mal entendido hasta lo bien argumentado. En el drama de Bernard Shaw titulado *El discípulo del diablo*, cuando Judith Anderson interpreta que la disposición de Richard Dudgeon de ser colgado en lugar de su esposo se debe a la simpatía que siente por él o al amor que siente por ella, Richard niega categóricamente: "Lo que hice anoche lo hice a sangre fría, sin preocuparme por tu marido, o por ti, ni la mitad de lo que me preocupo por mí mismo. No tuve ningún motivo y ningún interés: sólo puedo decirte que cuando me pregunté si sacaría mi cuello del lazo y pondría allí el de otro hombre, supe que no podría hacerlo".[21]

La característica del compromiso que me interesa más aquí es el hecho de que mete una cuña entre la elección personal y el bienestar personal, y gran parte de la teoría económica tradicional se basa en la identidad de ambos. Esta identidad se ve oscurecida a veces por la ambigüedad del término

[21] G. B. Shaw, *Three Plays for Puritans* (Harmondsworth, 1966), p. 94.

"preferencia", ya que el uso normal de la palabra permite la identificación de la preferencia con el concepto del mejoramiento, y al mismo tiempo es natural que se defina lo "preferido" como lo "escogido". No tengo una postura dogmática sobre el uso "correcto" de la palabra "preferencia", y me sentiría satisfecho mientras no se hagan ambos usos *simultáneamente,* intentando una afirmación empírica por medio de dos definiciones.[22] La conexión básica entre el comportamiento de elección y el logro del bienestar en los modelos tradicionales se rompe en cuanto se admite el compromiso como un ingrediente de la elección.

V

"Muy bien —diría usted—, ¿pero qué tan pertinente es todo esto para la clase de elecciones que interesa a los economistas? La economía no tiene mucho que ver con la marcha de Richard Dudgeon hacia la horca." Me parece que debiéramos convenir de inmediato en que, para muchos tipos de comportamiento, es improbable que el compromiso sea un ingrediente importante. En la compra privada de muchos bienes de consumo, puede estar limitado el campo para el ejercicio del compromiso, y quizá se muestre raramente en actos exóticos tales como el boicoteo de los aguacates sudafricanos o la evitación de las festividades españolas. Por lo tanto, es posible que el compromiso no plantee

[22] Véase mi ensayo "Behaviour and the Concept of Preference", *Economica,* 40 (1973); y Schick, "Rationality and Sociality".

un problema importante en muchos estudios del comportamiento del consumidor y sus interpretaciones. Es posible que ni siquiera la simpatía sea demasiado importante, porque las fuentes de la interdependencia personal se encuentren en otra parte, por ejemplo en el deseo de igualarse a los vecinos o en la influencia de los hábitos de otras personas.[23]

Pero la economía no se ocupa sólo del comportamiento del consumidor; ni el consumo se limita a los "bienes privados". Un área donde cobra gran importancia la cuestión del compromiso es la de los llamados bienes públicos. Estos bienes deben contrastarse con los "bienes privados" que no pueden ser usados por más de una persona: si usted se come un pedazo de pastel de manzana, yo no podré hacerlo también. No ocurre así con los "bienes públicos", por ejemplo con una carretera o un parque público, que usted y yo podemos usar al mismo tiempo. En muchos modelos económicos sólo se consideran los bienes privados, y así ocurre típicamente cuando la "mano invisible" queda encargada de hacer el bien visible. En realidad, sin embargo, los bienes públicos son importantes en la mayoría de las economías y abarcan un gran conjunto de servicios, desde las carreteras y el alumbrado público hasta la defensa. Abunda la infor-

[23] Véase a J. S. Duesenberry, *Income, Saving and the Theory of Consumer Behavior* (Cambridge, Mass, 1949); S. J. Prais y H. S. Houthakker, *The Analysis of Family Budgets* (Cambridge, 1955); W. Gaertner, "A Dynamic Model of Interdependent Consumer Behaviour", mimeografiado (Universidad Bielefeld, 1973); R. A. Pollak, "Interdependent Preferences", *American Economics Review,* 66 (1976).

mación en el sentido de que la participación de los bienes públicos en el consumo nacional ha crecido marcadamente en la mayoría de los países del mundo.

También se ha discutido mucho el problema de la asignación óptima de bienes públicos, sobre todo en la bibliografía económica reciente.[24] En particular, se ha prestado gran atención al problema de la revelación correcta de las preferencias. Este problema es muy evidente en el caso de los esquemas de suscripción, donde se cobra a cada persona de acuerdo con los beneficios recibidos. El problema principal se centra en el hecho de que a todos conviene minimizar el beneficio esperado, pero esto puede conducir al rechazo de un proyecto público que se habría justificado si se conocieran los beneficios verdaderos. El análisis de esta dificultad, conocida a veces como el problema del "viaje gratis", ha producido recientemente algunas propuestas muy ingeniosas para superar esta ineficiencia en el marco de la acción egoísta.[25] Se

[24] Véase E. Lindahl, *Die Gerechtigkeit der Besteuerung* (Lund, 1919), traducido al inglés en R. A. Musgrave y A. Peacock, *Classics in the Theory of Public Finance* (Londres, 1967); P. A. Samuelson, "The Pure Theory of Public Expenditure", *Review of Economic Studies,* 21 (1954); R. Musgrave, *The Theory of Public Finance* (Nueva York, 1959); L. Johansen, *Public Economics* (Amsterdam, 1966); D. K. Foley, "Lindahl's Solution and the Core of an Economy with Public Goods", *Econometrica,* 38 (1970); E. Malinvaud, "Prices for Individual Consumption, Quantity Indicators for Collective Consumption", *Review of Economic Studies,* 39 (1972).

[25] T. Groves y J. Ledyard, "Optimal Allocation of Public Goods: A Solution to the 'Free Rider Problem' ", Ensayo para Discusión núm. 144 (Centro de Estudios Matemáticos

establece el mecanismo de remuneración con tal astucia diabólica que los individuos se ven motivados a revelar exactamente su verdadera disposición a pagar por el bien público en cuestión. Una dificultad de esta solución deriva de una supuesta limitación de las posibilidades estratégicas al alcance del individuo, cuya eliminación conduce a un resultado imposible.[26] Otra dificultad se refiere al hecho de que al dar a los individuos el incentivo para revelar la verdad se entrega dinero, de modo que la distribución del ingreso se modifica sin control de las consideraciones distributivas. Por supuesto, este efecto puede eliminarse mediante una redistribución de las dotaciones iniciales y las participaciones en los beneficios,[27] pero

en Economía y Ciencia de la Administración, Universidad Northwestern, 1975); J. Green y J. J. Laffont, "On the Revelation of Preference for Public Goods", Informe Técnico núm. 140 (Instituto de Estudios Matemáticos en las Ciencias Sociales, Universidad de Stanford, 1974). Véase también J. Dreze y D. de la Vallee Poussin, "A Tatonnement Process for Public Goods", *Review of Economic Studies,* 38 (1971); E. Malinvaud, "A Planning Approach to the Public Goods Problem", *Swedish Journal of Economics,* 73 (1971); V. L. Smith, "Incentive Compatible Experimental Processes for the Provision of Public Goods", mimeografiado (Reunión de Verano de la Sociedad Econométrica, Madison, 1976).

[26] Véase J. Ledyard y D. J. Roberts, "On the Incentive Problem for Public Goods". Ensayo para Discusión núm. 116 (CMSEMS, Universidad Northwestern, 1974). Véase también L. Hurwicz, "On Informationally Decentralized Systems", en R. Radner y B. McGuire, *Decisions and Organizations* (Amsterdam, 1972).

[27] Véase el Teorema 4.2 de Groves y Ledyard, "Optimal Allocation of Public Goods".

tal acción plantea obviamente sus propias dificultades.

En este problema es fundamental el supuesto de que, cuando se le pregunta algo, el individuo ofrece la respuesta que maximizará su ganancia personal. ¿Hasta dónde es correcto este supuesto? Dudo que sea muy bueno en general. ("¿Dónde está la estación del ferrocarril?", me pregunta. "Allí", contesto, señalando la oficina de correos, "¿y de paso podría echar esta carta al buzón?" "Sí", contesta el otro, decidido a abrirla para ver si contiene algo valioso.) Aun en el contexto particular de la revelación de preferencias en materia de bienes públicos, es posible que el comportamiento maximizador de las ganancias no sea el mejor supuesto. Leif Johansen, uno de los autores que más han aportado a la economía pública, tiene razón cuando cuestiona el supuesto en este contexto:

> En este, como en otros campos, la teoría económica tiende a sugerir que los individuos son honestos sólo en la medida en que tengan incentivos económicos para serlo. Éste es un supuesto del *homo economicus* que dista mucho de ser evidente, y que debe ser confrontado con las realidades observadas. En efecto, una línea de pensamiento simple sugiere que el supuesto no puede ser cierto en su forma más extrema. Ninguna sociedad sería viable sin ciertas normas y reglas de conducta. Tales normas y reglas son necesarias para la viabilidad exactamente en los campos donde están ausentes los incentivos estrictamente económicos y no pueden crearse.[28]

[28] L. Johansen, "The Theory of Public Goods: Misplaced Emphasis" (Instituto de Economía, Universidad de Oslo, 1976). Véase también J. J. Laffont, "Macroeconomic Constraints, Economic Efficiency and Ethics", mimeografiado (Universidad de Harvard, 1974); P. Bohm, "Estimating Demand for Public Goods: An Experiment", *European Economic Review*, 3 (1972).

No se trata de saber si los individuos ofrecen invariablemente una respuesta honesta a cada pregunta, sino si ofrecen siempre una respuesta maximizadora de la ganancia, o por lo menos lo hacen con tanta frecuencia que este supuesto general resulte apropiado para la teoría económica. La presencia de respuestas que no maximizan las ganancias, incluidas las verdaderas, revela de inmediato la presencia del compromiso como parte del comportamiento.

Esta cuestión es pertinente también para la bibliografía reciente sobre la votación estratégica. Recientemente se han establecido varios resultados analíticos excelentes que demuestran la imposibilidad de que cualquier procedimiento de votación satisfaga ciertos requerimientos elementales y haga de la votación honesta la estrategia maximizadora de la ganancia para todos.[29] No se cuestiona la

[29] A. Gibbard, "Manipulation of Voting Schemes: A General Result", *Econometrica*, 41 (1973); M. A. Satterthwaite, "Strategy-proofness and Arrow's Conditions", *Journal of Economic Theory*, 10 (1975); D. Schmeidler y H. Sonnenschein, "The Possibility of Non-manipulable Social Choice Functions" (CMSEMS, Universidad Northwestern, 1974); B. Dutta y P. K. Pattanaik, "On Nicely Consistent Voting Systems" (Escuela de Economía de Delhi, 1975); P. K. Pattanaik, "Strategic Voting without Collusion under Binary and Democratic Group Decision Rules", *Review of Economic Studies*, 42 (1975); B. Peleg, "Consistent Voting Systems" (Instituto de Matemáticas, Universidad Hebrea, Jerusalén, 1976); A. Gibbard, "Social Decision, Strategic Behavior, and Best Outcomes: An Imposibility Result"

corrección de estos resultados, ¿pero es correcto el supuesto de que los individuos tratan siempre de maximizar sus ganancias personales en su comportamiento de votación? En las elecciones grandes, resulta difícil demostrar que cada votante tiene alguna esperanza real de afectar el resultado con su voto, y si la votación implica algún costo, la ganancia neta esperada de la votación puede ser típicamente negativa. Sin embargo, todavía puede ser muy elevada la proporción de votos en grandes elecciones, y en otra parte he tratado de demostrar que, en tales elecciones, los individuos pueden estar guiados a menudo "no tanto por la maximización de la utilidad esperada como por algo mucho más simple, o sea el deseo de registrar su verdadera preferencia".[30] Si este deseo refleja un sentimiento de compromiso, el comportamiento en cuestión no correspondería a la visión del hombre en la teoría económica tradicional.

VI

La cuestión del compromiso es importante en varios otros contextos económicos.[31] Es fundamental

Ensayo para Discusión núm. 224 (CMSEMS, Universidad Northwestern, 1976).

[30] Véase A. K. Sen, *Collective Choice and Social Welfare* (Edimburgo y San Francisco, 1970), p. 195.

[31] Véase la exposición que presenta Ragnar Frisch de la necesidad de "un fundamento teórico realista para la política social", en su "Samarbeid mellon Politikere og økonometrikere om Formuleringen av Politiske Preferenenser" (Socialøkonomen, 1971). (Estoy en deuda con Leif Johansen, quien tradujo para mí las porciones pertinentes de este ensayo.) Véase también J. A. Mirrlees, "The Economics

para el problema de la motivación del trabajo, cuya importancia para la producción no puede pasarse por alto.

Es sin duda costosa, y quizá imposible, la elaboración de un sistema de supervisión con recompensas y castigos de tal manera que todos tengan incentivos para esforzarse. Por lo tanto, todos los sistemas económicos han tendido a depender de la existencia de actitudes hacia el trabajo que se colocan por encima del cálculo de la ganancia neta derivada de cada unidad de esfuerzo. El acondicionamiento social desempeña aquí un papel muy importante.[32] Estoy persuadido de que las actuales dificultades económicas de Gran Bretaña tienen mucho que ver con problemas de la motivación para el trabajo que caen fuera de la economía de recompensas y castigos, y una de las razones de que los economistas parezcan tener tan poco que aportar en esta área es la omisión, en la teoría económica tradicional, de todo este tema del compromiso y las relaciones sociales que lo rodean.[33]

of Charitable Contributions", Reunión de la Sociedad Econométrica Europea (Oslo, 1973).

[32] Véase A. Fox, *Beyond Contract: Work, Power and Trust Relations* (Londres, 1974); H. G. Nutzinger, "The Firm as a Social Institution: The Failure of a Contractarian Viewpoint", Ensayo de Trabajo núm. 52 (Instituto Alfred Weber, Universidad de Heidelberg, 1976).

[33] "Tampoco... debemos olvidar la medida en que la teoría convencional pasa por alto la forma y la razón por la que se organiza el trabajo dentro de la empresa y el establecimiento, lo que podríamos llamar las 'relaciones sociales', del proceso de producción", R. A. Gordon, "Rigor and Relevance in a Changing Institutional Setting", Discurso Presidencial, *American Economic Review*, 66 (1976).

Por supuesto, estas cuestiones se conectan con la ética, ya que el razonamiento moral influye sobre nuestras acciones, pero en un sentido más amplio se trata de cuestiones culturales, entre las que se incluye la moral. En efecto, para mencionar un caso extremo, uno de los objetivos primordiales de la "revolución cultural" china fue el incremento del sentido de compromiso teniendo en mente los resultados económicos: "el objetivo de la Gran Revolución Cultural Proletaria es una revolución en la ideología del pueblo, para alcanzar resultados mayores, más rápidos, mejores y más económicos en todos los campos del trabajo".[34] Por supuesto, China estaba experimentando con una reducción extraordinaria del papel de los incentivos materiales en la producción, lo que seguramente habría incrementado el papel asignado al compromiso, pero aun dentro de los sistemas de pagos tradicionales, suele recurrirse en gran medida a reglas de conducta y modos de comportamiento que van más

Véase también R. Dahrendorf, *Class and Class Conflict in Industrial Society* (Stanford, 1959); O. E. Williamson, "The Evolution of Hierarchy: An Essay on the Organization of Work", Ensayo para Discusión núm. 91 (Universidad de Pensilvania, 1976); y S. A. Marglin, "What Do Bosses Do? The Origins and Functions of Hierarchy in Capitalist Production", *Review of Radical Political Economics*, 6 (1974).

[34] "The Decision of the Central Committee of the Chinese Communist Party Concerning the Great Proletarian Cultural Revolution", adoptada el 8 de agosto de 1966, reproducida en Joan Robinson, *The Cultural Revolution in China* (Harmondsworth, 1969). Véase también A. K. Sen, *On Economic Inequality* (Oxford, 1973); y C. Riskin, "Maoism and Motivation: A Discussion of Work Motivation in China", *Bulletin of Concerned Asian Scholars*, 1973.

allá de los incentivos estrictamente económicos.[35] Cuando se administra una organización *enteramente* mediante incentivos de ganancia personal se está condenado al fracaso.

Más adelante veremos lo que podría encontrarse detrás del sentido de compromiso. Ahora quisiera subrayar que la moral o la cultura en que se basa tal sentimiento podría ser más bien estrecha, muy lejana de la grandeza de enfoques tales como el utilitarismo. Las "colusiones implícitas" que se han observado en el comportamiento empresarial de los oligopolios parecen basarse en un sistema de confianza mutua y un sentido de responsabilidad que tiene límites bien definidos, de modo que los intentos de "universalización" de la misma clase de comportamiento a otras esferas de la acción pueden resultarle extraños. Se trata de manera precisa de una cuestión de ética empresarial que se supone aplicable en un campo estrictamente limitado.

De igual modo, en las negociaciones salariales y los contratos colectivos, puede tener límites bien definidos el sentimiento de solidaridad en ambos bandos, y quizá no encaje nada bien en un enfoque tal como el del utilitarismo general. El supuesto implícito de Edgeworth antes mencionado, en el sentido de que el egoísmo y el utilitarismo agotan las posibilidades de motivaciones alternativas, será especialmente estéril en este contexto. Es posible que el campo del compromiso sea extenso, y que el compromiso basado en el utilitarismo y

[35] Véase Williamson, "The Evolution of Hierarchy", donde aparece un análisis crítico de la literatura reciente sobre esta área.

otros sistemas morales universalizados constituya una parte relativamente pequeña del conjunto total.

VII

A veces se critica la teoría económica de la utilidad —que se relaciona con la teoría del comportamiento racional— por ser demasiado estructurada; se alega que los seres humanos son "más simples" en la realidad. Si nuestro argumento es correcto hasta aquí, ocurrirá precisamente lo contrario: la teoría tradicional tiene *muy escasa* estructura. Se asigna *un* ordenamiento de preferencias a una persona, y cuando es necesario se supone que este ordenamiento refleja sus intereses, representa su bienestar, resume su idea de lo que debiera hacerse, y describe sus elecciones y su comportamiento efectivo. ¿Podrá hacer todo eso un ordenamiento de preferencias? Una persona así descrita puede ser "racional" en el sentido limitado de que no revele inconsistencias en su comportamiento de elección, pero si no puede utilizar estas distinciones entre conceptos muy diferentes, diremos que es un tonto. En efecto, el hombre *puramente* económico es casi un retrasado mental desde el punto de vista social. La teoría económica se ha ocupado mucho de este tonto racional arrellanado en la comodidad de su ordenamiento *único* de preferencias para todos los propósitos. Necesitamos una estructura más compleja para acomodar los diversoss conceptos relacionados con su comportamiento.

¿Qué clase de estructura necesitamos? La importante distinción de John Harsanyi, entre las preferencias "éticas" de una persona y sus preferencias "subjetivas", abre un campo más amplio: "las primeras deben expresar lo que este individuo prefiere (o preferiría, mejor dicho) de acuerdo con consideraciones sociales impersonales exclusivamente, y las últimas deben expresar lo que en efecto prefiere, sobre la base de sus intereses personales o cualquier otra base".[36] Esta estructura doble nos permite distinguir entre lo que considera bueno una persona desde el punto de vista social y lo que considera bueno desde su propio punto de vista. Presumiblemente, la simpatía interviene en forma directa en la llamada preferencia subjetiva, pero el papel del compromiso queda poco claro. En la medida en que se cree que las preferencias "subjetivas" de una persona "definen su función de utilidad", parece existir la intención de excluir de allí el compromiso, pero surge una ambigüedad del hecho de que las preferencias subjetivas "expresan sus preferencias plenamente, tal como son en realidad". ¿Se entenderá esto en el sentido de la elección, o en el sentido de la concepción de su propio bienestar? Es posible que Harsanyi haya querido referirse a esto último, porque a las preferencias "éticas" se asigna por contraste el papel de expresar "lo que prefiere el individuo sólo en los momentos posiblemente raros en los que se impone una actitud imparcial e impersonal especial".[37] ¿Pero qué diremos si el individuo se desvía

[36] J. Harsanyi, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy*, 63 (1955), p. 315.

[37] *Ibid.*, pp. 315-316.

de su maximización del bienestar personal (incluida cualquier simpatía), no por efecto de una preocupación imparcial por todos,[38] sino por un sentimiento de compromiso hacia algún grupo particular, digamos el de los vecinos o el de la clase social a la que pertenece? La verdad es que todavía nos falta estructura.

Aun en su expresión de juicios morales desde un punto de vista impersonal, resulta deficiente una estructura *doble*. Un ordenamiento de preferencias puede ser *más* ético que otro, pero *menos* que un tercero. También en este sentido necesitamos más estructura. He propuesto en otra parte —en la conferencia de Bristol de 1972 sobre "la razón práctica"— que debemos considerar *ordenamientos de ordenamientos de preferencias* para expresar nuestros juicios morales.[39] Quisiera examinar un poco más esta estructura. Puede verse una

[38] Adviértase que, según Harsanyi, "las preferencias de un individuo satisfacen este requerimiento de impersonalidad si indican la situación social que escogería si no supiera cuál sería su posición general en la nueva situación escogida (y en cualquiera de sus opciones), sino que tuviera una probabilidad igual de obtener cualquiera de las posiciones sociales existentes en esta situación, desde la más alta hasta la más baja" (p. 316).

[39] A. K. Sen, "Choice, Orderings and Morality", en S. Körner (comp.), *Practical Reason* (Oxford, 1974). Véase también la contestación de J. Watkins y mi réplica en el mismo volumen, y R. C. Jeffrey, "Preferences among Preferences", *Journal of Philosophy*, 71 (1974); K. Binmore, "An Example in Group Preference", *Journal of Economic Theory*, 10 (1975), y B. A. Weisbrod, "Toward a State-Preference Model of Utility Function Preferences: A Conceptual Note", mimeografiado (Universidad de Wisconsin, 1976).

moral particular, no sólo en términos del ordenamiento "más moral" del conjunto de acciones alternativas, sino como un ordenamiento moral de los ordenamientos de las acciones (algo que va mucho más allá de la mera identificación del ordenamiento "más moral" de las acciones). Sea X el conjunto de combinaciones alternativas y mutuamente excluyentes de las acciones consideradas, y sea Y el conjunto de ordenamientos de los elementos de X. Llamaremos un metaordenamiento del conjunto de acciones X a un ordenamiento del conjunto Y (integrado por ordenamientos de acciones). Sostengo que un ordenamiento particular del conjunto de acciones X no está suficientemente articulado para decir mucho acerca de una moral dada, y que se obtiene un formato más robusto escogiendo un metaordenamiento de acciones (es decir, un ordenamiento de Y, en lugar de X). Por supuesto, tal metaordenamiento puede incluir entre otras cosas la especificación de un ordenamiento particular de acciones como el "más moral", pero en la medida en que el comportamiento efectivo pueda basarse en una transacción entre las exigencias de la moral y la búsqueda de varios otros objetivos (incluido el interés propio), debemos tomar en consideración también las posiciones morales relativas de los ordenamientos de acciones que *no* son "más morales".

Por ejemplo, consideremos un conjunto X de combinaciones alternativas de acciones y los tres ordenamientos siguientes de este conjunto de acciones X: el ordenamiento A representa mi bienestar personal (o sea que en cierto sentido representa mis intereses personales); el ordenamiento B re-

fleja mis intereses personales "aislados", omitiendo la simpatía (cuando tal separación es posible, lo que no siempre ocurre),[40] y el ordenamiento *C* en cuyos términos realizo efectivamente mis elecciones (cuando tales elecciones pueden representarse mediante un ordenamiento, lo que tampoco ocurre siempre).[41] El ordenamiento "más moral", *M*, puede ser concebiblemente cualquiera de estos ordenamientos *A, B* o *C*. O puede ser algún otro ordenamiento muy diferente de estos tres. (Así ocurrirá si las elecciones de acciones efectivas no son las "más morales" en términos del sistema moral en cuestión, y si además el sistema moral requiere el sacrificio de algún interés propio y también del interés propio "aislado".) Pero aun cuando se identifique algún ordenamiento *M*, distinto de *A, B* y *C*, como el ocupante del sitio más alto de la tabla moral, no sabremos cuál es el lugar ocupado por *A, B* y *C*. Por ejemplo, si se coloca la búsqueda del interés propio, con inclusión del placer y el dolor derivados de la simpatía, por encima de la búsqueda del interés propio "aislado" (lo que conduce a una coincidencia parcial del interés propio con la moral), y las elecciones efectivas reflejan una posición moralmente superior a la búsqueda del interés propio (quizá debido a un compromiso en la dirección moral), la

[40] Esto presupone cierta "independencia" entre los diversos elementos que influyen sobre el nivel del bienestar general, lo que implica cierta "posibilidad de separación" Véase W. M. Gorman, "Tricks with Utility Functions", en M. Artis y A. R. Nobay (comps.), *Essays in Economic Analysis* (Cambridge, 1975).

[41] Véase la nota 16 anterior.

moral en cuestión precipitará el metaordenamiento *M, C. A, B,* en orden descendente. Por supuesto, esto va mucho más allá de la especificación de que *M* es "moralmente lo mejor".

La técnica del metaordenamiento permite una extensión variable de la articulación moral. No se está sosteniendo que un metaordenamiento moral deba ser un ordenamiento *completo* del conjunto *Y*, es decir, que deba ordenar por completo todos los ordenamientos de *X*. Puede ser un ordenamiento *parcial*, y espero que a menudo sea incompleto, pero creo que en la mayoría de los casos no habrá problema para ir mucho más allá de la expresión limitada permitida por la especificación doble de las preferencias "éticas" y "subjetivas".

Por supuesto, los ordenamientos de la acción pueden ordenarse también sobre bases distintas de las de un sistema de moral particular: el metaordenamiento es una técnica general utilizable bajo interpretaciones alternativas de la relación de metaordenamiento. Puede utilizarse para describir una ideología particular o un conjunto de prioridades políticas o un sistema de intereses de clase. En un contexto muy diferente, puede proveer el formato para expresar las preferencias que habríamos preferido tener ("Quisiera que me gustaran más los alimentos vegetarianos", o bien "Quisiera que no me gustara tanto fumar"). O puede usarse para analizar los conflictos implicados en la adicción ("Dados mis gustos actuales, me siento mejor con la heroína pero la heroína me lleva a la adicción, y habría preferido no tener estos gustos"). La herramienta de los metaordenamientos puede usar-

se en muchas formas diferentes en contextos distintos.

Claramente, ésta no es la ocasión propicia para un análisis detallado de la forma en que esta estructura más amplia permite un entendimiento mejor de la preferencia y el comportamiento. Desde luego, una estructura no es una teoría, y pueden formularse teorías alternativas utilizando esta estructura. Sin embargo, debo mencionar que la estructura exige mucho más información que la provista por la observación de las elecciones efectivas de los individuos, las que a lo sumo revelarían sólo el ordenamiento *C*. La estructura asigna un papel a la introspección y la comunicación. Para ilustrar un uso del aparato, mencionaré algunos resultados técnicos. Supongamos que estoy tratando de investigar su concepción de su propio bienestar. Usted especifica primero el ordenamiento *A* que representa su ordenamiento del bienestar. Pero quiero ir más allá y tener una idea de su función de utilidad *cardinal,* es decir, en términos generales, no sólo cuál ordenamiento le da a usted mayor bienestar, sino también en qué medida. Ahora le pido que ordene los diversos ordenamientos en términos de su "proximidad" a su ordenamiento efectivo del bienestar *A,* así como un policía usa la técnica del retrato hablado: ¿se parece esto más a él, o esto otro? Si sus respuestas reflejan el hecho de que la reversión de una preferencia más fuerte aleja más el resultado que la reversión de una intensidad menor de la preferencia, sus respuestas satisfarán ciertas propiedades de consistencia, y la sucesión de los ordenamientos nos permitirá comparar sus *diferencias* de bienestar entre pares. En efecto, considerando ordenamientos cada vez mayores, podremos determinar su función de bienestar cardinal con tanta precisión como usted quiera especificar.[42] No estoy afirmando que este tipo de diálogo sea el mejor método para descubrir su función de bienestar, pero ilustra que una vez que abandonamos el supuesto de que la observación de las elecciones es la única fuente de los datos sobre el bienestar, se abre todo un mundo nuevo, liberándonos del yugo de información del enfoque tradicional.

Esta estructura más amplia tiene muchos otros usos, por ejemplo el de permitir un análisis más claro de la *akrasia* —la debilidad de la voluntad— y aclarar algunas consideraciones conflictivas de la teoría de la libertad, los que he tratado de examinar en otra parte.[43] También ayuda a analizar

[42] Este resultado, y otros relacionados, surgieron en discusiones que sostuve con Ken Binmore en 1975; sin embargo, no hemos escrito el ensayo conjunto que proyectamos para publicar estos resultados. R. Nader-Ispahani trabaja ahora también sobre este punto.

[43] Véase Sen, "Choice, Orderings and Morality"; y también Sen, "Liberty, Unanimity and Rights", *Economica*, 43 (1976). Adviértase también la pertinencia de esta estructura para el análisis del carácter incompleto de la concepción de la libertad en términos de la capacidad para hacer *lo que en realidad queremos hacer.* "Si descubro que puedo hacer poco o nada de lo que deseo, sólo tengo que contraer o extinguir mis deseos, y estaré libre. Si el tirano (o 'persuasor oculto') logra condicionar a sus súbditos (o clientes) para que pierdan sus deseos originales y abracen ('interioricen') la forma de vida que ha inventado para ellos, habrá logrado liberarlos de acuerdo con esta definición." I. Berlin, "Two Concepts of Liberty", en *For Essays on Liberty* (Oxford, 1969), pp. 139-140.

el desarrollo del comportamiento que implica un compromiso en situaciones caracterizadas por juegos tales como el del Dilema de los Prisioneros.[44] Este juego se trata a menudo, con cierta justicia, como el caso clásico del fracaso de la racionalidad individualista. Hay dos jugadores y cada uno tiene dos estrategias, las que podemos llamar egoísta y altruista para facilitar la memoria sin tener que entrar en gran detalle. Cada jugador se encuentra mejor personalmente jugando la estrategia egoísta, independientemente de lo que haga el otro, pero ambos estarán mejor si escogen la estrategia altruista en lugar de la egoísta. Individualmente es óptima la conducta egoísta: sólo podemos afectar nuestra propia acción y no la del otro, y dada la estrategia del otro —cualquiera que sea— cada jugador estará mejor siendo egoísta. Pero esta combinación de estrategias egoístas, derivada del egoísmo de ambos, produce un resultado peor para ambos que el resultado de la estrategia altruista escogida por ambos. Puede demostrarse que este conflicto existirá no importa que el juego se repita muchas veces.

Algunas personas consideran desconcertante el hecho de que el egoísmo individual produzca un resultado inferior para todos, pero éste es sin duda un conflicto bien conocido, discutido en términos generales durante largo tiempo. En efecto, fue la

[44] Véase R. D. Luce y H. Raiffa, *Games and Decisions* (Nueva York, 1958); A. Rapoport y A. M. Chammah, *Prisoners' Dilemma: A Study in Conflict and Cooperation* (Ann Arbor, 1965); W. G. Runciman y A. K. Sen, "Games, Justice and the General Will", *Mind*, 74 (1965); N. Howard, *Paradoxes of Rationality* (Cambridge, Mass, 1971).

base de la famosa distinción de Rousseau entre la "voluntad general" y la "voluntad de todos".[45] Pero el enigma desde el punto de vista del comportamiento racional reside en el hecho de que, en las situaciones reales, los individuos no siguen a menudo la estrategia egoísta. Son bien conocidos los ejemplos de la vida real de este tipo de comportamiento en circunstancias complejas, pero aun en experimentos controlados en las condiciones del laboratorio, los individuos que juegan el Dilema de los Prisioneros se comportan frecuentemente de modo altruista.[46]

Al interpretar estos resultados experimentales, el teórico de los juegos se ve tentado a imputarlos a la falta de inteligencia de los jugadores: "Evidentemente, los jugadores ordinarios no son lo bastante refinados en términos estratégicos para entender que la estrategia DD [la estrategia egoísta] es la única defendible en términos racionales, y esta deficiencia intelectual les impide el cierre".[47] Es posible que un enfoque más fecundo permita la posibilidad de que la persona sea *más* refinada que lo establecido por la teoría, que se pregunte cuál tipo de preferencia le gustaría que tuviera el otro jugador y, sobre bases kantianas, haya considerado el caso en que él mismo tuviera esas preferencias, o se comportara *como si* las tuviera. Esta línea de razonamiento le obliga a considerar las

[45] Véase Runciman y Sen.

[46] Véase, por ejemplo, L. B. Lave, "An Empirical Approach to the Prisoner's Dilemma Game". *Quarterly Journal of Economics*, 76 (1962), y Rapoport y Chammah, *Prisoners' Dilemma*.

[47] Rapoport y Chammah, p. 29.

modificaciones del juego que se generarían si se actuara mediante el compromiso (en términos de las "preferencias reveladas", esto parecería *como si* tuviera preferencias diferentes de las que en efecto tiene), y tuviera que evaluar las normas de comportamiento alternativas en esas condiciones. He explorado estos problemas en otra parte,[48] de modo que aquí señalaré simplemente que el aparato del *ordenamiento de ordenamientos* ayuda al razonamento que implica la consideración de los méritos de diversos tipos de preferencias (o del hecho de actuar como si se tuvieran).

VIII

Por supuesto, la admisión del comportamiento basado en el compromiso tendría consecuencias importantes para la naturaleza de muchos modelos económicos. He tratado de mostrar por qué es necesario este cambio y por qué pueden ser graves las consecuencias. Quedan pendientes muchas cuestiones, incluida la importancia empírica del compromiso como parte del comportamiento, la que variará de un campo a otro. He indicado también que la información empírica necesaria no puede buscarse en la mera observación de las elecciones efectivas, y debe abarcar otras fuentes de información, la introspección y la discusión.

[48] Sen, "Choice, Orderings and Morality". Véase también K. Baier, "Rationality and Morality", y A. K. Sen, "Rationality and Morality: A Reply", próxima publicación en *Erkenntnis;* K. Baier, *The Moral Point of View* (Ithaca, 1958); y el análisis de Fred Schick, "Rationality and Sociality".

Sin embargo, subsiste el problema de determinar si esta concepción del hombre equivale a verlo como una criatura irracional. Mucho depende del concepto de racionalidad que se use, y existen muchas caracterizaciones alternativas. En el sentido de la *consistencia* de la elección, no hay razón para pensar que la admisión del compromiso deba implicar alguna desviación de la racionalidad. Pero éste es un sentido débil de la racionalidad.

El otro concepto de la racionalidad predominante en la economía la identifica con la posibilidad de justificar cada acto en términos del interés propio: cuando el acto x es escogido por la persona i y el acto y es rechazado, esto implica que i espera que sus intereses personales resulten mejor servidos por x que por y. Me parece que hay tres elementos distintos en este enfoque. Primero, es una visión de consecuencias: se juzgan los actos sólo por las consecuencias.[49] Segundo, es un enfoque de evaluación de los *actos,* antes que de evaluación de las *reglas.* Y tercero, las únicas consecuencias consideradas en la evaluación de los actos son las que afectan nuestros propios intereses, todo lo demás es a lo sumo un producto intermedio. Desde luego, pueden disputarse las pretensiones de cada uno de estos elementos de ser una parte necesaria de la concepción de la racionalidad en el sentido del diccionario de "el poder de ser capaz de ejercer la razón propia". Además, no es difícil encontrar ar-

[49] Sobre la naturaleza del "consecuencialismo" y los problemas engendrados por este concepto, véase B. Williams, "A Critique of utilitarianism", en J. J. C. Smart y B. Williams, *Utilitarianism: For and Against* (Cambridge, 1973).

gumentos para rechazar la camisa de fuerza de cada uno de estos tres principios. La justificación de las acciones basadas en el compromiso puede derivar de la violación de cualquiera de estos tres principios. El compromiso se relaciona a veces con un sentimiento de obligación que va más allá de las consecuencias. A veces se acepta la ausencia de ganancia personal en *actos* particulares considerando el valor de las *reglas* de comportamiento. Pero aun dentro de un marco de evaluación de los actos por sus consecuencias, la exclusión de toda consideración distinta del interés propio parece imponer una limitación totalmente arbitraria a la noción de racionalidad.

Henry Sidgwick señaló la naturaleza arbitraria del supuesto del egoísmo:

> Si el Utilitarista debe contestar este interrogante: "¿Por qué debo sacrificar mi propia felicidad en aras de la mayor felicidad de otro?", sin duda será admisible que se pregunte al Egoísta: "¿Por qué debo sacrificar un placer presente en aras de un placer futuro? ¿Por qué debo preocuparme por mis propios sentimientos futuros más que por los sentimientos de otras personas?" Sin duda parecerá paradójico para el Sentido Común que busquemos una razón para buscar nuestra propia felicidad en conjunto; pero no entiendo cómo puedan repudiar por absurda esta demanda quienes adoptan las concepciones de la escuela empírica extrema de psicología, aunque se supone de ordinario que tales concepciones tienen una afinidad estrecha con el Hedonismo Egoísta. Es cierto que el Ego es sólo un sistema de fenómenos coherentes, que el "Yo" permanente e idéntico no es un hecho sino una ficción, como sostienen Hume y sus seguidores; entonces ¿por qué habría de preocuparse una parte de la serie de sentimientos en la que se resuelve el Ego por otra parte de la misma serie, en mayor medida que con cualquiera otra serie?[50]

La concepción de la racionalidad que la identifica con una evaluación de las consecuencias de los actos puede cuestionarse desde cualquiera de estos tres ángulos. La admisión del compromiso como parte del comportamiento no implica negar la evaluación razonada como base de la acción.

No tiene caso dedicar grandes esfuerzos a debatir la definición "apropiada" de racionalidad. El término se emplea en muchos sentidos diferentes, y ninguna de las críticas de los fundamentos conductistas de la teoría económica presentadas aquí depende de la definición escogida. La cuestión principal es la aceptabilidad del supuesto de la búsqueda invariable del interés propio en cada acto. Cuando decimos que ese tipo de comportamiento es racional, o que sus desviaciones son irracionales, no cambiamos la pertinencia de estas críticas, aunque producimos una definición arbitrariamente estrecha de la racionalidad. Este ensayo no se ha ocupado en el tema de si el comportamiento humano se describe mejor como racional o irracional. La tesis principal ha sido la necesidad de incluir el com-

[50] H. Sidgwick, *The Method of Ethics* (Londres, 1874; 7ª ed., 1907), pp. 418-419. Véase también la vigorosa exposición que hace Nagel de la tesis de que "el propio altruismo depende de un reconocimiento de la realidad de otras personas, y de la capacidad equivalente para considerarse simplemente como un individuo entre muchos", *The Possibility of Altruism*, p. 1.

promiso como parte del comportamiento. El compromiso no presupone el razonamiento, pero tampoco lo excluye; en efecto, en la medida en que las consecuencias para los demás deban entenderse y evaluarse con mayor claridad en términos de nuestros valores e instintos, es posible que se amplíe el campo del razonamiento. He tratado de analizar las extensiones estructurales en la concepción de la preferencia que vuelve necesario el comportamiento basado en la evaluación razonada del compromiso. Las preferencias entre ordenamientos deben ser sustituidas por una estructura más rica que implica los metaordenamientos y los conceptos relacionados.

También he sostenido que no debe verse el comportamiento en términos de la dicotomía tradicional existente entre el egoísmo y los sistemas morales universalizados (como el utilitarismo). Los grupos intermedios entre el individuo y los demás, como la clase y la comunidad, proveen el foco de muchas acciones que implican el compromiso. Por lo tanto, el rechazo del egoísmo como una descripción de la motivación no implica la aceptación de cierta moral universalizada como la base del comportamiento efectivo. Tampoco vuelve a los seres humanos excesivamente nobles.

Y por supuesto, el uso del razonamiento tampoco implica una sabiduría notable.

Es tan cierto como que el nombre de César era Flavio,
Que ningún economista fue jamás más sabio,

dijo Robert Frost en una alabanza jocosa del economista contemporáneo. Es posible que deba rendirse un tributo dudoso semejante al hombre económico de nuestra concepción modificada. Si este hombre brilla en absoluto, brillará por comparación —en contraste— con la imagen dominante del tonto racional.