

# Influencia de la psicología milliana en la teoría económica: problemática asociada al caso particular del comportamiento altruista

José Luis HERRANZ GUILLÉN  
Departamento de Estudios y Proyectos  
del Organismo Autónomo Agencia Local de Empleo y Formación.  
Ayuntamiento de Getafe  
[joseluis.herranz@alefgetafe.org](mailto:joseluis.herranz@alefgetafe.org)

**Received:** 29/12/2015

**Accepted:** 22/02/2016

## Resumen:

El altruismo es una experiencia y un tipo de comportamiento universales que han sido estudiados por los economistas. Sin embargo, los supuestos normalmente admitidos sobre los fundamentos psicológicos de la naturaleza humana por las doctrinas utilitaristas, como es el caso de la Economía, han causado diversos problemas que dificultan la formulación de una teoría general del altruismo. Esto se debe principalmente a que el altruismo utilitarista no es realmente altruismo. Explicamos la teoría psicológica que une las doctrinas utilitarias de James y John Stuart Mill, y mostramos cómo sus principales contenidos, y problemas teóricos en relación con el altruismo, aparecen nuevamente replicados en los modernos modelos neoclásicos, como el de Gary S. Becker, que explican el altruismo con un análisis formal más desarrollado. Proponemos la necesidad de actualizar la tradición utilitarista sobre el altruismo, considerando los actuales descubrimientos de la neurociencia cognitiva. Curiosamente, estos hallazgos conectan el funcionamiento del cerebro con los supuestos teóricos considerados por Adam Smith, y la tradición de la psicología de la simpatía.

**Palabras clave:** Altruismo, Utilitarismo, Intersubjetividad, James Mill, John S. Mill, Gary S. Becker

## Abstract:

Altruism is an universal experience and kind of behavior that has been studied by economists. However, the assumptions frequently accepted about psychological foundations of human nature in utilitarian doctrines, as it's the case of Economics, have caused several problems that make difficult to formulate a general theory of altruism. That is mainly because utilitarian altruism is not really altruism. We explain the psychological theory that link James & John S. Mill utilitarian doctrines, and show how it's principal contents, and theoretical problems dealt with altruism, appear again replicated in modern neoclassical models, as Gary S. Becker's, that explain altruism by means of more developed formal analysis. We propose the need for updating utilitarian tradition on altruism, taking into account the current findings in cognitive neuroscience. Curiously, these findings connect brain running with the theoretical assumptions considered by Adam Smith, and the psychology of sympathy tradition.

**Key-words:** Altruism, Utilitarianism, Intersubjectivity, James Mill, John S. Mill, Gary S. Becker

**JEL Classification:** D64; D01

## 1. Introducción

Shotland y Stebbins (1983, p. 36) señalan que la literatura sobre el altruismo se enmarca dentro de dos tradiciones intelectuales, que a su vez se plasman en dos enfoques teóricos no integrados. Uno consiste en admitir la existencia de un genuino altruismo no intencionalmente utilitario, por el que los individuos experimentan el impulso -sea de origen innato o adquirido- de ayudar a otras personas soportando el coste de dicha ayuda como recursos “perdidos”. El segundo enfoque no contempla la acción de ayuda como un “gasto” a fondo perdido, sino como una inversión utilitaria. En este caso el altruismo viene condicionado por la expectativa de un beneficio neto -ya se trate de una gratificación psicológica de bienestar, o de cualquier beneficio material, presente o futuro- derivado de la acción altruista.

La mayor parte de las teorías que se valen del supuesto de un egoísmo universal aceptan la motivación única del egoísmo, y explican la conducta altruista como resultado indirecto de un

primario interés egoísta (Elster 2000 [1979], p. 243). Esta tradición puede ilustrarse con la tesis de Michael Ghiselin en su obra *The economy of nature and the evolution of sex* (1974), cuando desacredita la hipótesis de un altruismo genuino en la naturaleza afirmando irónicamente que si se araña a un (presunto) altruista se puede comprobar cómo sangra un hipócrita (real) (op. cit., p. 247). O sea, el altruismo viene a ser la excepción que confirma la regla del egoísmo.

Aunque existe una larga y amplia tradición de enfoques utilitaristas, si intentamos definir la línea que los une podemos afirmar que se trata de doctrinas que asumen el axioma del egoísmo motivacional: toda acción humana, según sus defensores, está motivada por un propósito acumulador y posesivo -crematístico, hedonista, eudemonista, etc.- de utilidad. Las teorías basadas en el utilitarismo suelen apoyarse igualmente en el supuesto del solipsismo, doctrina según la cual ya se trate de un solipsismo radical (sólo existe el yo-sujeto), o bien de una versión más relajada (sólo el yo y su conciencia son objetos de conocimiento verificable), lo que se asume es una limitación epistemológica inherente en lo que llamamos intersubjetividad. Esta limitación viene a ser un refuerzo y una causa del egoísmo, dado que una mente solipsista está abocada a tomar decisiones motivadas en lo único que puede conocer: su propia experiencia.

Las reservas expresadas a esta tradición pueden ilustrarse con la crítica de Nagel en el sentido de que el egoísmo es un prejuicio teórico basado en una idea errónea: “puesto que soy yo quien actúa, incluso cuando actúo en interés de otro debe haber un interés mío que proporcione el impulso” (Nagel 2004 [1970], p. 90). Obviamente, que el yo resulte motivado por emociones que “parten del organismo” no implica necesariamente que esa fuerza interior sea conductual o psicológicamente egoísta: “la creencia de que un acto mío beneficiará a alguien más puede motivarme sólo porque quiero su bien o quiero algo que lo implica” (ibíd., p. 91).

La psicología actual acepta la tesis nageliana, que puede reconocerse en la experiencia mental de integración intersubjetiva: “sensación de ‘pertenecer a’ algo, asociándonos a ello o ‘dejándonos arrastrar’ por ello” (“empatía”, en *The Oxford Companion to Mind* editado por Gregory 2004). Asimismo ocurre respecto a la entrada “simpatía” en *The Dictionary of Psychology* editado por Corsini (1999), caracterizada como “apreciación sensible o preocupación emocional por, y el acto de compartir el estado emocional y mental de otra persona, o animal, o de un grupo”.

De otro lado, y alineado en buena medida con lo anterior, el solipsismo representa en palabras de Hannah Arendt la falacia filosófica más persistente y posiblemente perniciosa que haya existido (Arendt 2002 [1978], p. 71). El arquetipo ideal de sujeto solipsista, la *res cogitans* cartesiana, “no podría saber que existe algo como la realidad, ni sospechar la posible distinción entre lo real y lo irreal, entre el mundo común de la vigilia y el no-mundo privado de los sueños” (ibíd., p. 73).

En este trabajo se estudian las líneas maestras de la psicología que expone James Mill en su obra *Analysis of the phenomena of the human mind* (1829), y que su hijo John Stuart discutió y complementó en diversos trabajos para formar su concepto de utilidad. Se analiza cómo, aunque sin una continuidad estrictamente lineal, en ambos casos se trata de una visión egoísta y solipsista de la naturaleza humana que se encuentra presente también en la posterior ciencia económica.

Que la teoría económica moderna está construida sobre la base de una racionalidad identificada con el egoísmo psicológico es un hecho admitido explícita o implícitamente por la mayoría de los economistas (Frank 1994, p. 227). Que la mente humana es solipsista y no puede traspasar la barrera psicológica del ego, ya fue escueta y contundentemente planteado por Jevons: “Cada mente es inescrutable para cualquier otra, y no es posible ningún denominador común de sentimientos” (Jevons 1998 [1871], p. 75). Esta tesis ha pervivido y desembocado en los postulados psicológicos de influyentes economistas del s. XX, como Robbins:

(...) cuando hago comparaciones interpersonales (como por ejemplo, cuando estoy decidiendo entre reclamaciones que afectan a la satisfacción de dos niños enredados en una riña infantil), mis juicios son más bien juicios de valor que juicios acerca de hechos verificables. (Robbins, 1938, p. 640).

Las críticas del utilitarismo como doctrina, y al concepto de utilidad de la teoría económica estándar, son numerosas y diversas en la literatura especializada (por ejemplo Trincado 2009 para una crítica filosófica, y Frey y Stutzer 2002 para el caso de la economía de la felicidad). Sin embargo, no se ha profundizado suficientemente en la idea de que los postulados psicológicos en que se sustentan las teorías utilitaristas son la consecuencia de una teoría de la mente problemática, incapacitada de raíz para explicar una tipología de comportamientos muy relevante en la especie humana, como es el caso del altruismo. La teoría psicológica que soporta el utilitarismo de los Mill ya se topó con el problema del altruismo, lo cual no fue obstáculo para su propagación a la teoría económica moderna a través de múltiples versiones de la función de utilidad que comparten una concepción egoísta y solipsista del sujeto. Un caso destacado por su importancia e impacto en la literatura es el intento de Gary S. Becker de introducir el altruismo en la teoría de neoclásica de la utilidad. Becker propone un altruismo de interdependencias utilitarias, motivado por la conveniencia e inferencia analógica de los estados mentales ajenos. Tal concepción del altruismo, bastante discutida por especialistas de la economía (p.ej. Rutherford 2010), debe ser comprendida además como un producto derivado de las líneas de fondo en que se sustenta la teoría psicológica “milliana”. Este producto difiere claramente de la concepción basada en la simpatía e identificación con las experiencias psicológicas de otros yoes.

En el artículo se analizan los principales problemas filosóficos y teóricos del paradigma utilitarista milliano en su explicación del comportamiento altruista, así como de la teoría económica del altruismo inspirada en el modelo general de Becker. Se explica cómo en ambos casos se acusan las mismas deficiencias que parten de un concepto erróneo del altruismo, el cual emana necesariamente de una teoría de la mente y de la intersubjetividad incorrectas. Ambas ya están superadas por la psicología y las ciencias cognitivas actuales, y curiosamente estas coinciden en mayor medida con la que llamamos vía no explorada por los Mill: la de la psicología de la simpatía cuya referencia obligada es la obra psicológica de Adam Smith.

## **2. El comportamiento altruista como problema teórico en la Economía**

El altruismo es un tipo empírico de comportamiento presente en las relaciones humanas, y existe abundancia de registros que permiten mostrar una “arqueología de la compasión” de tiempo muy pretérito en la especie (Spikins et al. 2010). Desde la óptica meramente conductista observamos que, sea cual sea la motivación psicológica “interna”, el altruismo se manifiesta en que los individuos ayudan “externamente” (revelan la preferencia de hacerlo) a otras personas, hacen entregas de bienes y servicios sin contraprestación, atienden a otros en apuros, e incluso llegan a sacrificar de manera importante, a veces, su riqueza, libertad, salud y la propia vida en favor de otras personas, causas sociales, intelectuales y morales, y compromisos con sistemas de reglas y creencias -por ejemplo religiosas, filosóficas, científicas o políticas-. Esto es así ya se trate de personas vinculadas por lazos de consanguinidad o no, o de culturas coincidentes con la propia o por el contrario ampliamente distanciadas. Además, practicando el comportamiento comprometido (con reglas de conducta), se transfieren recursos a otras personas en circunstancias no estrictamente relacionadas con las conductas de ayuda o de refuerzo -a través del comportamiento- de una determinada identidad personal o social.<sup>1</sup> Las acciones de ayuda -como atender a alguien herido en la calle o dar una limosna a un indigente-, de regalo -dar una propina en un restaurante, o los típicos arreglos amistosos de la vida cotidiana- y de compromiso o limitación autoimpuesta -como renunciar al bienestar presente en favor de otros, siguiendo intencionalmente una determinada regla de comportamiento asumida-, tienen en común que el agente sacrifica sus recursos “aquí y ahora” sin recibir a cambio una compensación cierta y claramente definida -por ejemplo a través de un pacto informal o contrato formal- ni presente ni futura. Puede opinarse que quien sacrifica algo por otros lo hace *siempre* porque busca de forma pre-

---

<sup>1</sup> Valga la definición de Sen (1977, p. 327) sobre el compromiso: “una persona que elige actuar de una manera que cree que le producirá un nivel de bienestar personal inferior a otra alternativa también disponible para ella. Téngase en cuenta que la comparación se hace entre niveles de bienestar *esperados*, y por lo tanto esta definición del compromiso excluye los actos contrarios al propio interés resultantes de un error de pronóstico sobre las consecuencias” (cursivas en el original).

meditada algún tipo de compensación y pretende obtener una ganancia neta por ello, como en la famosa crítica de Ghiselin y la doctrina utilitarista inspiradora de la teoría económica. Pero este argumento es un prejuicio heredado de una antigua e influyente tradición filosófica.<sup>2</sup> Como explica Boulding:

Resulta tentador para el economista argumentar que en realidad no hay regalos y que todas las transacciones contienen algún tipo de intercambio, o sea, alguna clase de *quid pro quo*. Si depositamos una limosna en el platillo de un ciego, lo hacemos porque el ciego nos da algo a cambio. Sentimos cierta satisfacción emocional de virtud, y eso es lo que recibimos por nuestra limosna (...). Incluso si miramos la transacción como un intercambio, no obstante, se trata claramente de uno muy curioso. La limosna en el platillo es una transferencia de activos del donante al beneficiario suficientemente clara. Sin embargo, lo que circula desde el beneficiario al donante es algo misterioso. (Boulding 1971, p. 236).

Las conductas de altruismo y compromiso pueden caracterizarse como anomalía económica (Thaler 1992), dado que aunque en ocasiones están asociadas a una previa intención utilitaria, como conseguir una recompensa material futura o un bienestar psicológico presente; sin embargo no siempre está claro que ese sea el estímulo motivador. Ayudar, regalar y renunciar a algo valioso en favor de otros son comportamientos objetivamente altruistas y no siempre es posible para el observador evaluar y cuantificar qué parte del propósito que conduce a la acción es subjetivamente altruista y qué parte es egoísta.<sup>3</sup> El altruismo y el egoísmo han sido una fuente de importantes polémicas filosóficas a lo largo de la historia y sólo en fechas relativamente recientes -apenas cuatro o cinco décadas- han aparecido teorías científicas, aún en fase de revisión, para explicar este tipo de conductas que, en palabras que aún resuenan de Edward O. Wilson, constituyen “el problema teórico central” de la biología evolutiva (Wilson 1980, p. 3).

¿Es el altruismo un problema teórico central también para la economía? En apariencia la corriente académica más extendida de la teoría económica moderna no ha encontrado dificultad explicativa intrínseca, aportando diversidad de líneas de investigación (Fontaine 2007). Por su parte, el utilitarismo de John Stuart Mill ha sido sometido a multiplicidad de interpretaciones en las que incluso cabe admitir el comportamiento motivado por expresiones de la utilidad tales como la felicidad, el anhelo del bien común, y la conducta moral en general (Brown 1973; Crisp 1997). Sin embargo, la cooperación y su caso extremo, el altruismo, siguen planteando en la actualidad contradicciones y desafíos teóricos que invitan a considerar estas conductas como problema teórico en la economía.<sup>4</sup> Ello es así porque la economía es una ciencia basada en el artificio teórico de la maximización de utilidad, es decir, construida sobre el supuesto de una psicología egoísta aunque estemos explicando comportamientos morales o “altruistas”. Por otra parte, el individualismo metodológico heredado de la tradición utilitarista plantea limitaciones explicativas de la interacción social empírica, en la que hay una notable cantidad de conocimiento social y altruistamente construido y transmitido, y abundantes comportamientos advertida o inadvertidamente pro sociales (Arrow 1994; Samuelson 1993). Puede argumentarse que el axioma del egoísmo utilitario y la metodología individualista constriñen las posibilidades de explicar esa parte de la acción humana que consiste en “dar” sin hacerlo con la intención de recibir un beneficio mayor (De Juan y Monsalve 2006, p. 106). Comencemos presentando los antecedentes filosóficos doctrinalmente más influyentes en la teoría económica moderna, para

<sup>2</sup> Véase al respecto Searle (2000). Blaug (1985, p. 203) opina que el egoísmo motivacional es una de esas proposiciones metafísicas ampliamente extendidas que resultan prácticamente irrefutables.

<sup>3</sup> Actualmente sí es posible conocer la intención subyacente a la preferencia altruista revelada a través de la tecnología y experimentación neurocientíficas (Lamm et al. 2007). El comportamiento humano ha dejado de ser una caja negra motivacional típica de la tradición psicológica conductista.

<sup>4</sup> Elster (2000 [1979], pp. 235 y ss.) presenta el altruismo como uno de los casos de racionalidad problemática, en el sentido de difícil de explicar desde las premisas de la teoría de la elección racional en la que se basa, por otra parte, la teoría económica moderna.

explicar a continuación con un ejemplo destacado cómo ha integrado la teoría neoclásica el comportamiento altruista en su concepción utilitaria de la elección.

### 3. Psicología milliana del sujeto y la intersubjetividad

La teoría psicológica en la que se fraguó la economía política decimonónica está basada en la tradición principal del empirismo filosófico que conecta el pensamiento de autores como Hobbes–Locke–Hume, sistematizado en la visión recogida por Bentham (Watson 1966). La herramienta explicativa clave es la teoría asociacionista de las experiencias psicológicas, complementada con el principio utilitarista (véanse Sober y Wilson 2000, Introducción; y De Waal 2006, cap. 1). Esta teoría sostiene una perspectiva hedonista de la motivación, pero además, se trata de una perspectiva individualista y egocéntrica, en el sentido de que se basa en postulados solipsistas dentro del subjetivismo radical que ha caracterizado tanto a la filosofía racionalista continental como al empirismo británico (García Suárez 1976, p. 111). Uno de los trabajos que mejor representa este enfoque es la obra de James Mill *Analysis of the phenomena of the human mind* (1829), destacado colaborador y discípulo de Bentham. Su hijo John Stuart, conocedor de las teorías psicológicas de vanguardia en su época, como el evolucionismo spenceriano, y amigo del prestigioso psicólogo asociacionista Alexander Bain, reeditó póstumamente el libro de su padre en 1869, pocos años antes de morir, lo cual informa de que las ideas expuestas por James Mill en esta publicación todavía disfrutaban de interés y buena acogida en la segunda mitad del siglo XIX. Mill hijo, por su parte, aunque discrepó abiertamente del dogmatismo utilitarista de Bentham y su padre; sin embargo no se separó sustancialmente de la teoría psicológica de estos, una doctrina normativa, egocéntrica y abiertamente utilitarista.

James Mill reduce en su *Analysis* los fenómenos mentales a sus unidades básicas: los sentimientos que aparecen a la conciencia (*feelings*) y que son las sensaciones y las ideas, para después explicar la vida mental de todo individuo como asociaciones de esas unidades básicas. Dichas asociaciones generan un orden psíquico al que se puede llamar “realidad”. Avanzando el discurso milliano hallamos, como resultara previsible en un seguidor de Bentham, la que podemos llamar perspectiva hedonista presente en la obra: el individuo *busca* con sus acciones, y *así ha de hacerlo*, su propio placer neto de dolor (una expresión del concepto “propio interés”).

La aritmética del placer y del dolor permite concebir otra unidad de análisis, fundamental para explicar la motivación, que es la utilidad. La utilidad es una experiencia que contribuye a configurar el sujeto experimentador y “propietario” de experiencias utilitarias. En efecto, para esta línea de pensamiento asociacionista existe una mente consciente que se construye a través de la sedimentación de una identidad psicológica en el espacio-tiempo. La identidad personal, como en Locke, es la consciencia del acontecer, como si las experiencias conocidas se adhirieran en “algo” preexistente, un observador identificado, formando un bucle cognoscitivo que, por el hecho de ser conocido y experimentado en primera persona pasa a ser una especie de propiedad personal de acceso privado. El orden dentro del bucle lo aporta la memoria y esta se halla sometida a la categoría del tiempo. Así pues, el yo se concibe como “el recuerdo de una cadena de estados de conciencia” (Mill 2001[1829], pp. II y 132), o sea, una identidad histórico–narrativa.

Procede traer ahora a colación que en la teoría psicológica milliana sólo es posible conocer las “propias” experiencias subjetivas si bien, en un intento de esquivar la poderosa atracción del solipsismo los Mill -padre e hijo- se valen de la capacidad mental de inferencia analógica para explicar las más que evidentes, y prácticas, ideas y sentimientos que se forma todo individuo acerca de la experiencia mental de los otros yoes. Esta teoría, no obstante, forma parte del acervo teórico de la tradición solipsista, aunque no se trate de la versión más radical. En efecto, según este punto de vista los individuos no estamos epistemológicamente capacitados para acceder al conocimiento directo de la experiencia que ocurre en otras mentes. Nuestra intersubjetividad está limitada de raíz y, por ello, la información con que se cuenta para interaccionar con otros agentes se halla naturalmente restringida -obsérvese la carga escéptica de este corolario filosófico-psicológico-. El modo de conocer la experiencia del otro es mediante la simulación analógica basada en la experiencia propia, una especie de hipótesis o “copia” simulada, o “recreación” mental en primera persona de lo que somos capaces de inferir que ocurre en otra persona a la que observamos o imaginamos:

El conocimiento de mis propios estados de conciencia es conciencia en sí misma, del momento presente, y el recuerdo de esa conciencia en todo el pasado. De los estados de conciencia de otros hombres no tengo conocimiento directo. Yo dibujo mi creencia en ellos solo mediante las señales que recibo. Esas señales sólo son significativas por referencia a mis propios estados de conciencia (...) (en consecuencia) es necesario tener en cuenta la diferencia entre la idea anexa al pronombre *yo* y aquella otra anexa al pronombre *tú*. (Mill 2001[1829], vol. II, p.135; paréntesis añadido, cursivas en el original).

John Stuart defendió un argumento prácticamente idéntico en su ensayo *An examination of Sir William Hamilton's philosophy* (1865), donde plantea una peculiar teoría del realismo filosófico basada en presentar el mundo y la materia como posibilidades permanentes de sensación que la mente asocia generando una realidad ordenada y coherente.<sup>5</sup> En efecto, si somos capaces de experimentar ese orden en referencia a las cosas que nos rodean, entonces podemos trazar inferencias racionales basadas en la inducción (la experiencia en primera persona) para “atisbar” qué experiencias es más probable que estén viviendo otros yoes. Se trata de una toma de conciencia puramente imaginaria, dado que según el subjetivismo radical las experiencias ajenas son incognoscibles. No obstante, la inferencia nos permite esquivar el escepticismo absoluto salvando a la vez la perspectiva egocéntrica y, en cierto modo, flirtear con el idealismo sin arriesgar el axioma de que existe un mundo objetivo aunque, como apuntara Locke en su momento y posteriormente recalcará magistralmente Hume, la inferencia por analogía no produce certeza, sino sólo probabilidad. La cita relevante aquí es la siguiente reflexión de John Stuart:

Concluyo que otros seres humanos tienen sentimientos como yo porque, en primer lugar, tienen cuerpos como el mío lo cual sé, en mi propio caso, que es la condición antecedente de los sentimientos; y porque, en segundo lugar, exhiben los actos, y los signos externos que en mi propio caso sé por experiencia que son causados por sentimientos. Soy consciente en mí mismo de una serie de hechos conectados por una secuencia uniforme, de la cual el comienzo es la modificación de mi cuerpo, el medio son los sentimientos, el final es el proceder exterior. En el caso de los demás seres humanos tengo la evidencia de mis sentidos para el primero y el último eslabones de la serie, pero no para el eslabón intermedio. (Mill 1979[1865], p. 191).

La teoría de la mente como caja negra, que será el punto de apoyo de la psicología conductista del siglo posterior así como de la propia teoría económica neoclásica -tan influida por el conductismo imperante en buena parte del s. XX-, queda trazada en el argumento expuesto de John Stuart Mill, que amplía del siguiente modo:

Encuentro, sin embargo, que la secuencia entre el primero y el último es tan regular y constante en estos otros casos como lo es en el mío. En mi propio caso sé que el primer eslabón produce el último a través del eslabón intermedio, y que no podría producirlo sin él. La experiencia, por tanto, me obliga a concluir que debe haber un eslabón intermedio; el cual debe, o bien ser el mismo en los demás que en mí mismo, o bien diferente; debo, o bien creerlos vivos, o bien creerlos autómatas; y al creerlos vivos, es decir, al suponer que el eslabón es de la misma naturaleza que en el caso del cual tengo experiencia, y que es similar en todos los demás respectos, someto a los demás seres humanos, en cuanto a fenómenos, bajo las mismas generalizaciones que sé por experiencia que son la verdadera teoría de mi propia existencia. (Mill 1979[1865], p. 191).

Llegados a este punto, es preciso resaltar el papel desempeñado por el esfuerzo cognitivo de la voluntad, tal y como queda expuesto por Mill hijo, lo cual hace de la operación de inferencia un proceso singularmente racional, consciente y premeditado. Esto distancia y hasta aparta la teoría de la intersubjetividad milliana de la moderna teoría de la mente (*ToM*) propia de la neurociencia cognitiva actual, que defiende la suposición innata e inconsciente en la especie *Homo sapiens* de que sus individuos disponen de estados mentales que causan sus intenciones y compor-

<sup>5</sup> Semejante y tal vez inspirada en ella es la teoría de la percepción que el filósofo analítico John L. Austin recuperó, ya en el s. XX, en su libro *Sense and sensibilia* (1959).

tamientos de acuerdo a reglas universales. De este modo, los organismos dotados de teoría de la mente, caso de los niños y niñas con más de tres años de edad, infieren espontáneamente y sin esfuerzo cognitivo, a través de procesos psicológicos inconscientes, esos estados mentales de otros a través de supuestos, interrelaciones lógicas, y la observación de signos externos que señalizan información, de lo cual emerge un conocimiento intersubjetivo que, sin dejar de ser hipotético, se le atribuye un alto grado de certeza adaptada a cada circunstancia (Saxe y Kanwisher 2003).

En otro orden de cosas, la concepción hedonista de la naturaleza humana dentro de esta tradición filosófica es un asunto importante en la explicación de la alteridad mediante una particular teoría utilitarista de la cooperación. Y es que James Mill (y antes Hobbes, Hume y Bentham) nos da la clave posible para hacer al *homo œconomicus* un agente egoísta, sí, pero capaz de juicio moral y de cooperación, incluso en su sentido más fuerte de altruismo. Aunque la principal motivación placentera del género humano, dice Mill padre, es conseguir la riqueza, el poder y la dignidad individual -y evitar la pobreza, la impotencia y el desprecio de los demás-.<sup>6</sup> sin embargo también está provisto de afectividad, y eso es la causa de que podamos experimentar placer cuando observamos el bien (placer) de los demás, y dolor cuando observamos el mal (dolor) ajeno: “contemplamos a nuestro prójimo como causa de nuestros placeres, ya sea individualmente, o en grupo” (Mill 2001 [1829] vol. II, p. 173). Así, cuando trata sobre la bondad (*kindness*) escribe que: “No hay nada que asocie más rápidamente con las ideas de nuestro placer, y dolor, que la idea de los placeres y dolores de nuestros seres cercanos” (p. 175). Pero, una vez reconocida esta sintonía y correlación que se produce entre las emociones propias y ajenas, el autor vuelve a la teoría del solipsismo analógico: “nuestra verdadera idea de los placeres y dolores de otro hombre es solamente la idea de nuestros propios dolores, o de nuestros propios placeres, asociados con la idea de otro hombre” (p. 175). A esta conexión y resonancia hedónica entre las experiencias propias y ajenas James Mill la llama compasión (*compassion*), de modo que cuando el individuo actúa cooperando con los otros es porque esos otros forman parte de su propio interés, “le son propias” sus experiencias inferidas de placer y de dolor, y actúa de ese modo con el objeto de acrecentar las primeras y reducir las segundas porque eso le concierne en términos de “su propia utilidad”. Por ejemplo, los padres se sacrifican altruistamente por sus hijos porque ellos son una fuente de placer que James Mill denomina “asociación simpatizante” (*agreeable association*). De este modo, y a diferencia de otras explicaciones de la cooperación que circulaban en la época con carácter algo menos utilitarista, como la de James Mackintosh, la de Mill padre es completamente hedonista y consecuencialista.<sup>7</sup>

En efecto, el utilitarismo radical de James Mill se puede comprobar en el ataque que propinó a Sir James Mackintosh, en *Fragment on Mackintosh* (1835), a propósito de la publicación por parte de este de unos escritos sobre ética en la edición de 1829 de la *Enciclopedia Británica*. Mackintosh, que flirtea con la tesis escocesa del sentido moral (y por eso la aguda -e injusta- crítica de Mill) se alineaba con el utilitarismo benthamita en la aceptación del principio de utili-

---

<sup>6</sup> Según Mill las principales causas remotas del placer y del dolor son las relacionadas con tales circunstancias. Véase Mill (2001[1829] vol II, pp. 164 y ss.)

<sup>7</sup> Una crítica al consecuencialismo dogmático de Mill padre es la de su hijo John Stuart. En efecto, en un comentario a la afirmación de su padre de que nuestra idea de los placeres y dolores del otro no es más que la idea de nuestros propios placeres y dolores asociados con la idea del otro -postulado genuinamente solipsista- dice que es una proposición ambigua. Tal y como apostilla Copleston (2011, p. VIII-20 n. 40) aludiendo a la crítica de John Stuart a su padre: “Una cosa es decir que si yo me complazco en el placer de otro, el placer que yo siento es mío y no del otro (...). Pero algo distinto es deducir de ahí que si busco el placer del otro, lo hago como medio para conseguir mi propio placer”. A lo que está apuntando John Stuart en la crítica a su padre no es a la refutación del solipsismo -ambos mantuvieron un enfoque solipsista- sino a la refutación del hedonismo: es posible actuar altruistamente sin un móvil hedonista, lo cual lleva implícito que en la mente humana existe una conciencia que va más allá de la función de utilidad individualista-hedonista. Pero John Stuart no da el siguiente paso consistente en admitir el conocimiento directo -y motivador- de la experiencia ajena, ya que supondría una refutación del solipsismo en la que él no quiso implicarse: que podemos conocer, tener conciencia directa, de la función de utilidad de otros agentes de forma dissociada de la función de utilidad propia.

dad, y valiéndose del asociacionismo psicológico imperante defendió la moral como destinada a la felicidad general. En principio no había razones de peso para la reacción de Mill, pero este consideró que aquél había adulterado el principio de utilidad como explicación inmanente de la moralidad al apreciar que Mackintosh abrazaba peligrosamente la teoría del sentido moral originaria de la Escuela Escocesa, y aproximarse a la idea del juicio estético/moral -igual que Kant en su *Crítica del Juicio* (1787)- como desencadenante de cogniciones y emociones morales de forma independiente a la utilidad. Es decir, se trata de aplicar el sentido de lo bueno del mismo modo que utilizamos el sentido de lo bello a la hora de motivar o emitir un juicio acerca de la moralidad de una acción independientemente del balance utilitario. Mackintosh admitía el principio de utilidad como regla de moralidad elemental, pero insistió además en la existencia de sentimientos morales *a priori* que se experimentan espontáneamente ante la imaginación o la contemplación de un acto humano. Estos sentimientos están vinculados con el sentido de la belleza:

(...) las cualidades morales de un hombre virtuoso sirven realmente en cuanto que contribuyen a la felicidad o bien común. Pero uno puede aprobarlas y admirarlas sin más referencia a la utilidad que cuando aprecia un hermoso cuadro (...) De igual modo, los sentimientos que experimentamos al contemplar las cualidades indeseables de un hombre malo, no tienen por qué encerrar una referencia a su falta de utilidad. (Copleston 2011, p. VIII-21).

No resulta extraño pues pensar que, como argumenta Copleston (2011, p. VIII-21), lo que irritó a James Mill, un dogmático proponente del benthamismo, fue el alineamiento con Saftesbury -aunque se tratase de un alineamiento secundario- cuyas teorías había rechazado explícitamente Bentham, y en su afán reduccionista consideró superfluo admitir otras causas de la moralidad por muy bien acompañadas, coherentes y coincidentes que fueran con el principio de utilidad, único juicio moral que los utilitaristas ortodoxos admiten -todos los demás para ellos eran superfluos-. Así, “cualquier intento, como el de Mackintosh, de reconciliar el utilitarismo con la ética intuicionista sencillamente le indignaba” (Copleston 2011, p. VIII-21) Los benthamitas consideraban que el principio de utilidad era la regla de moral más afinada y elemental que existía, la única admisible, siendo todas las demás difusas, contradictorias o incompletas.

Pero ni Mackintosh, en su tibio intento de justificar la presencia de lo que hoy día llamaríamos una “gramática moral” congénita (Hauser 2006, p. 1), ni Mill hijo plantearon una crítica lo suficientemente certera al afán utilitarista tan extendido entre los empiristas, filósofos, psicólogos y economistas del s. XIX en Gran Bretaña. La idea es la siguiente: si favorecer el placer o aliviar el dolor ajenos nos aporta un rédito utilitario, y una vez conocidas podemos llegar a habituarnos a tales experiencias, entonces cabe preguntarse por la medida de la prudencia de dicha deriva asociativa, o sea: ¿cómo podemos arriesgar tanta utilidad realizando actos altruistas si resulta imposible verificar los placeres y dolores ajenos al tratarse de experiencias privadas?

El origen de esta crítica, la supuesta privacidad absoluta de la vida mental, ha sido objeto de una elevada controversia en el campo de la filosofía del s. XX. Su artífice es Wittgenstein en el *Tractatus logico-philosophicus* (1921). En efecto, si el solipsismo plantea la privacidad de los estados de conciencia, entonces toda aproximación a la intersubjetividad ha de ser, como defendieron los Mill y antes Hume, una cuestión de conocimiento probable a través de una mente capaz de llevar a cabo simulaciones y representaciones de realidad virtual. Pero desde la teoría de la mente milliana esto sólo puede ser posible si admitimos dos hipótesis: (a) que el conocimiento adquiera veracidad a través de alguna suerte de orden externo lógico-mental, una realidad consensuada que viene a ser necesariamente una construcción social, como el lenguaje, el imaginario cultural y el conocimiento tácito que caracterizan toda sociedad. Esto destruye la posibilidad de un individualismo metodológico radical. O (b), que exista un orden interno *a priori*, es decir, intuiciones certeras basadas en procesos epistemológicos innatos previos a la experiencia, y esto destruye la posibilidad de un empirismo ortodoxo.

Nos hallamos pues ante problemas de importancia en el plano teórico y filosófico. Discutamos si esta crítica puede enmendarse tras analizar el desarrollo impulsado por uno de los autores más destacados de la escuela económica neoclásica.



#### 4. La perspectiva beckeriana

Un marco de referencia posible en que los modelos neoclásicos emplazan el comportamiento altruista es el de un equilibrio de funciones de utilidad entre altruistas -que se ocupan de la utilidad de los demás, incluidos los egoístas-, y egoístas -que sólo se ocupan de su propia utilidad-. Estos últimos, que Sober (1996, p. 155) caracteriza como «trotamundos que van por libre», se benefician de las donaciones y ayudas de los individuos altruistas sin contribuir a la internalización de esa externalidad positiva, salvo que no les quede más remedio porque hacerlo les resulte más rentable que no hacerlo.

Gary S. Becker es el economista más influyente que desde la perspectiva neoclásica ha tratado el estudio del comportamiento altruista en diversos trabajos académicos. Ha sido además uno de sus principales divulgadores desde la teoría económica moderna, dado su interés académico por el estudio de las interacciones sociales.<sup>8</sup> En esencia, el modelo general que plantea Becker, que recoge la herencia utilitarista-marginalista-neoclásica, incluye la utilidad de otras personas en la función de utilidad del individuo altruista, y lo hace con una ponderación positiva -el odio, la envidia y otras emociones antisociales tendrían signo negativo-.<sup>9</sup> De este modo, el altruismo se explica como experiencia vicaria de un individuo (el altruista), que computa la utilidad ajena como una resonancia que le afecta positivamente en *su propia utilidad*. Se trata, por lo tanto, de una externalidad utilitaria, o interdependencia entre funciones de utilidad, y el altruista tendrá incentivo a ayudar a otra persona porque la unidad marginal de la ayuda -o de entrega en el caso del regalo, o del sacrificio en favor de otros en el caso del compromiso- le reporta más utilidad que dedicar esos recursos a su propio disfrute. En definitiva, el altruista ayuda, hace donaciones o se sacrifica libremente por los demás porque eso le permite maximizar *su* utilidad. El modelo puede representarse de diversas maneras, y queda formalizado básicamente como sigue (Becker 1987, cap. 8):<sup>10</sup>

$$U_i = f [ x_i , \psi (U_j) ]$$

siendo  $U_i$  es la utilidad del individuo con preferencias altruistas y  $U_j = g[x_j]$  la función de utilidad del individuo  $j$  que es, o puede ser, objeto del altruismo de  $i$ ;  $x_j$  representa los recursos que disfruta  $j$  (entre los que se incluyen las aportaciones del individuo  $i$ ) y  $x_i$  representa los recursos de los que disfruta  $i$ .  $U_j$  es la utilidad de otro individuo cuyo bienestar afecta al individuo altruista y  $\psi(U_j)$  es la experiencia utilitaria vicaria que tiene el individuo  $i$  acerca de la utilidad de  $j$ . Este último aspecto, que a nuestro entender no ha sido suficientemente comentado, conecta el modelo con la tradición solipsista y por esa misma razón se aparta de la psicología cognitiva,

---

<sup>8</sup> Nos centramos en la versión beckeriana, formalizada, inspiradora e influyente, si bien no la única, dentro del debate económico del altruismo en el s. XX. Para una síntesis acerca del debate aludido véase Fontaine (2007).

<sup>9</sup> Una posible expresión lineal de esta función de utilidad genérica sería:  $U_i = x_i + \beta \cdot x_j$ , donde  $\beta$  representa el *coeficiente de simpatía afectiva* del famoso modelo de Edgeworth (2000 [1881], pp. 89-93). Este coeficiente también podría denominarse, dependiendo de qué se considere que explica la externalidad utilitaria, coeficiente de relación afectiva, coeficiente de distancia social, coeficiente de intersubjetividad, de compromiso, de comunidad, de relación interpersonal, de acoplamiento simpatético, de identificación, etc. Dicho parámetro modula la fuerza del propio interés utilitario y lo complementa con el interés por la utilidad de otros individuos. Para un análisis más detallado de la visión de Edgeworth véase Collard (1975).

<sup>10</sup> Se han producido múltiples desarrollos posteriores basados en el modelo de Becker, algunos bastante recientes (véase Sobel 2005), y todos ellos caracterizados por presentar el altruismo como preferencias interdependientes que producen externalidades cruzadas en la función de utilidad. Estas externalidades se representan mediante alguna formalización de  $\beta$ , el parámetro modulador edgeworthiano que significa el gradiente del “eco” o “resonancia” de la utilidad ajena en la utilidad del propio agente decisorio. En las últimas décadas la neuroeconomía y la economía experimental y del comportamiento, además de otras disciplinas académicas como la psicología social y la neurociencia cognitiva, han conseguido avances en cuanto la comprensión, fundamento y cuantificación de dicho parámetro.

que admite la existencia de procesos mentales compartidos entre sujetos. Esto, queremos destacarlo, no es un asunto baladí.

Efectivamente, la cuestión clave en este momento del análisis, y en lo que aquí nos ocupa, es que Becker, en una formalización que revela con claridad su alineamiento con la teoría psicológica milliana, no incluye la utilidad ajena ( $U_j$ ) directamente en la función de utilidad; y por eso, si estamos describiendo la mente de un altruista, no le queda más remedio que introducir explícitamente una función transformada [ $\psi(U_j)$ ] que consiste en una representación mental a efectos prácticos.<sup>11</sup> Por un lado este proceder conserva las esencias del individualismo metodológico ortodoxo: no tenemos acceso directo a la experiencia utilitaria en segunda persona, y mucho menos a la posibilidad de cuantificar las utilidades ajenas y realizar comparaciones interpersonales con un propósito distributivo de maximización social. Por otro lado, aunque Becker esté utilizando una mera representación que no deja de ser un conocimiento probable, su inserción a efectos prácticos está negando el axioma asumido por la tradición central de la teoría económica; es decir, para teorizar el comportamiento altruista a Becker no le queda más remedio que hacer equilibrios metodológicos y aproximarse a la posibilidad de llevar a cabo, de algún modo aunque sea incompleto, comparaciones interpersonales de utilidad.<sup>12</sup>

Dado que se trata de preferencias altruistas y que los recursos a disposición son bienes, entonces  $\partial U_i / \partial x_i > 0$ ,  $\partial U_j / \partial x_j > 0$ ,  $\partial U_i / \partial U_j > 0$ , luego  $\partial U_i / \partial x_j > 0$ . La conducta racional consistirá pues en maximizar  $U_i$  sujeta a la restricción presupuestaria que venga al caso -en el modelo de Becker se trata de una adición de rentas o riquezas del conjunto de individuos entre los que se produce la interacción cooperativa. El autor la denomina renta o riqueza social-.

La ecuación puede expresarse como una función de los bienes  $x_i$  y  $x_j$  que disfrutaban cada uno de los individuos, y dentro del modelo caben diversas sofisticaciones añadidas introduciendo los matices formales oportunos, como por ejemplo que la externalidad utilitaria que disfruta el altruista esté asociada o no a que este lleve a cabo personalmente la ayuda o sacrificio en favor del individuo  $j$  (Becker 1997 [1974], p. 196n<sub>34</sub>). Si se trata de una persona que mejora su utilidad cuando toma conciencia de la mejora de la utilidad de  $j$ , vengan de donde vengan los recursos que recibe  $j$ , Becker dice que se trata de un genuino altruista (GA). Por el contrario, si para que el individuo en cuestión mejore su utilidad al mejorar la de  $j$  es preciso que este reciba una mejora derivada de la acción específica de  $i$ , Becker califica a  $i$  de falso altruista (FA) y su comportamiento “falsamente caritativo”. Esto es así porque, según el autor, lo que le aporta utilidad no es el bienestar de  $j$  en sí mismo (altruismo genuino), sino que *él* sea el artífice de la mejora en el bienestar de  $j$ : estaríamos ante un filántropo paternalista, o un “adicto” a las donaciones y sacrificios por los demás o incluso, como subraya Becker, “el deseo de evitar el desprecio de los demás o de recibir el beneplácito social”, o la importancia “de la culpabilidad y de otras secuelas de experiencias anteriores.”<sup>13</sup>

Dado que en este modelo se trabaja con el supuesto de la mente como caja negra, la motivación raíz parece indescifrable, puesto que el agente revela directamente sus preferencias, no sus motivaciones -aunque pueden ser inferidas o comprendidas por diversos procesos cognitivos, o a través de técnicas de laboratorio, pero el modelo no contempla dichas posibilidades-; y conocer estas motivaciones es irrelevante, según el autor, para el análisis económico. Ambos casos, GA y FA, se representarían como sigue:

<sup>11</sup> Dice Blaug (1985, p. 2010) en este sentido que: “si ninguno de nosotros fuese capaz de predecir nada sobre el comportamiento de otras personas, la propia vida económica, y no sólo las teorías sobre la vida económica, resultarían inimaginables”.

<sup>12</sup> Harsanyi (1955) asumió la posibilidad fáctica de las comparaciones interpersonales de utilidad defendiendo que estas no son las únicas hipótesis empíricas -todos las hacemos en nuestra vida cotidiana- basadas en un postulado no empírico. Boulding (1971, p. 246) especifica al respecto que una teoría del altruismo construida sobre la base de la interdependencia de las utilidades de los agentes requiere asumir supuestos sobre mensurabilidad y comparabilidad de las utilidades, aunque esto resulte teóricamente controvertido.

<sup>13</sup> Las citas corresponden respectivamente a Becker (1997 [1974], p. 196), y a Becker (1997 [1992], p. 246).

$$(GA) U_i = f [ x_i , x_j ] ,$$

donde  $\partial U_i / \partial x_i > 0$ , y  $\partial U_i / \partial x_j > 0$

$$(FA) U_i = h [ x_i , x_j (= d_i) ]$$

donde  $d_i$  representa las donaciones o sacrificios que realiza directamente  $i$  en favor de  $j$ , tal que  $\partial U_i / \partial x_i > 0$ ,  $\partial U_i / \partial d_i > 0$ , pero en este caso quedaría indeterminado el sentido de  $\partial U_i / \partial x_j$ .

### 5. Algunos problemas teóricos por resolver

El modelo beckeriano y sus desarrollos posteriores han sido objeto de diferentes críticas, algunas procedentes del análisis metodológico dentro del propio núcleo duro de la teoría neoclásica. Este es el caso de Blaug, quien en una crítica que todavía hoy puede considerarse vigente apostilló sobre la obra donde Becker expone el grueso de su teoría del altruismo, el *Tratado sobre la familia* (1981): “ignora el ‘amor’” (Blaug 1985, p. 272 n. 27). En efecto, la emoción del amor no es solamente una experiencia agradable o desagradable que pueda expresarse a efectos formales con el denominador común de la utilidad, sino que el amor consiste en un estado de conciencia ampliada en el que se produce una disipación del yo egocentrado y “aislado del mundo”, lo cual enriquece la gama de emociones (y utilidades) experimentadas. Actualmente ya se ha demostrado que a ello contribuye la existencia de grupos especializados de células encefálicas denominadas neuronas espejo. El sistema de neuronas espejo sitúa a los individuos que interactúan en planos emocionales y cognitivos mutuamente experimentados, que se explican por el acontecer de estados neurológicos semejantes funcionando en paralelo. Cuantas mayor atención e identificación existan entre las experiencias psicológicas interactivas, como ocurre entre otras en la experiencia amorosa, mayor capacidad tendrán los individuos de “leerse” y compartir recíprocamente sus vivencias personales. En la literatura sobre la cognición a través de neuronas espejo se denomina a este fenómeno reflejo especular (Iacoboni 2009).

Aunque ni los Mill ni Becker pudieron conocer los actuales descubrimientos neurocientíficos relacionados con las neuronas espejo, sin embargo sí existía ya en su tiempo una larga tradición de pensamiento social, y económico, además en la órbita del liberalismo, para la cual: “el amor es un sentimiento que lleva la atención de la mente más allá de sí misma, y su objeto se da en sentido de relación con algún semejante”, escribió Adam Ferguson (1974[1767], p. 16) un siglo antes de que Mill hijo reflexionara acerca de los estados de conciencia en su ensayo sobre la filosofía de William Hamilton. En esta sintonía, un coetáneo de Ferguson, Adam Smith, con buen criterio y adelantándose casi dos siglos y medio a los hallazgos neurocientíficos contemporáneos, expuso en varias partes de la *Teoría de los sentimientos morales* (1759) cómo los márgenes del yo se ven involuntariamente afectados por la simpatía con los demás, uno de cuyos estímulos acrecentadores es la contemplación del sufrimiento ajeno y la necesidad subjetiva -y no sólo objetiva- de ayuda (véase por ejemplo Smith 1997 [1759], p. I, i).<sup>14</sup>

Tal y como reconoce la ciencia psicológica actual el amor es una experiencia intersubjetiva, relacional e integradora, que por su propia naturaleza requiere reconocer la existencia subjetiva -y no sólo objetual- de un tú de carne y hueso además de la de un yo (Bartels y Zeki 2004). Por esta razón los intereses de ese tú no pueden ser atendidos (solo y siempre) por motivaciones egoístas como las del cálculo utilitario *a priori* del modelo de Becker. Cuando se ama, la atención puesta en los otros y sus intereses, la “compañía de sentimientos” que dice Smith, es un “principio de la naturaleza humana” (Smith 1997 [1759], p. I, iii, 1. p. 113), o sea, un instinto previo a la deliberación y al cálculo de intereses que hace posible la integración de las utilidades

---

<sup>14</sup> Un reconocimiento de Smith como antecedente destacado de la actual Behavioral Economics puede leerse en Ashraf et al. (2005). Aunque hemos citado a Smith, también pueden ser mencionados como economistas proponentes de visiones de la psicología humana alternativas a la del *homo oeconomicus* convencional las de, por ejemplo, *Ye-Machine* de Alfred Marshall, y el *orden sensorial* de Hayek (véanse respectivamente Frantz 2005, cap. 4, y Hayek 2004 [1952]).

por identificación. Por eso, las experiencias del amor y la simpatía no pueden explicarse rigurosamente desde la perspectiva del egoísmo motivacional.

Si la percepción de indicios que hacen altamente probable la necesidad de ayuda por parte de otros invita a aprovechar una buena oportunidad de maximizar la utilidad, entonces estamos ante un proceso motivacional egoísta -ya se trate de un “GA” o “FA” beckerianos- que hoy día sabemos experimentalmente que se diferencia como tal del proceso identificatorio-simpatizante que describieron Ferguson y Smith. En efecto, la neurociencia ha descubierto que en el cerebro se activan circuitos diferentes cuando se actúa “altruistamente” por una motivación egoísta -como puede ser evitar la culpa o la vergüenza, buscar la reputación social, ensalzar el ego y la autoestima, o sacudirse el estrés emocional consecuencia del contagio de las emociones ajenas-, y cuando se actúa por un genuino interés por el bienestar de una persona necesitada (Lamm et al. 2007; Singer y Lamm 2009). El segundo caso es un altruismo psicológico, en el que la motivación raíz es el bienestar ajeno, mientras que el primero, sin embargo, es egoísmo psicológico, con una motivación central que parte de y se conduce hacia el bienestar propio (Sober y Wilson 2000).

Aquí estriba la diferencia importante que traza la neurociencia cognitiva actual entre la experiencia psicológica intersubjetiva basada en la simulación (teoría de la mente), y aquella otra basada en la identificación (simpatía). La segunda conlleva compartir el sufrimiento con quien lo sufre en primera persona, mientras que la primera supone sólo la conciencia del mismo.<sup>15</sup> Se ha demostrado que cuando se percibe directamente, o se imagina, el dolor o sufrimiento psíquico de otra persona, se co-activan automáticamente procesos neurales que producen la representación de una experiencia psicológica similar a la que está experimentando quien padece esa emoción desagradable en primera persona. Esto no conlleva la co-activación de todo el entramado neurológico de la experiencia de dolor. El observador no siente la localización ni la intensidad del *dolor físico*; por el contrario, sí conoce la naturaleza del *dolor emocional* que está sintiendo el otro (su dimensión afectiva), pues en su cerebro se activan tanto la red de procesamiento del dolor (*pain matrix*) como los circuitos corticales que operan en funciones de mentalización más relacionadas con la propuesta de los Mill y de Becker (Singer et al. 2004). Esta experiencia afectivo-cognitiva puede desencadenar procesos motivacionales de acciones altruistas (empatía afectiva o simpatía), si bien la intensidad del co-sentimiento y el resultado final de la vivencia empática dependen de variables y circunstancias que intervienen en la neuromodulación del proceso cognitivo-emotivo-operativo (Hein y Singer 2008; Vignemont y Singer 2006).

De este modo, la vía no explorada por los Mill ni por Becker, la de la simpatía, o identificación psíquica del sujeto con otros sujetos ya sean empíricos, imaginarios e incluso abstracciones representativas de conjuntos de individuos, puede resultar suficientemente motivadora para producir acciones de genuino altruismo, en su sentido de “no utilitario” (Tyler y Blader 2000). Un exclusivo altruismo utilitario de inferencia por analogía, como el de Becker, vuelve a acusar los mismos problemas ya comentados en el caso del altruismo en la versión milliana. En primer lugar, el conocimiento probable de la interioridad psíquica ajena no debería ser suficientemente motivador para asumir el coste de oportunidad personal que supone ayudar o ceder altruistamente recursos, salvo que se sea un adicto patológico a los actos altruistas. Ello es así por un argumento de tipo evolutivo: el individuo incapaz de discriminar la realidad sobre la autenticidad y el grado de necesidad que experimentan otros se encuentra, *cæteris paribus*, y especialmente si se trata del tipo FA de Becker, en clara desventaja adaptativa frente a oportunistas, o simplemente frente a individuos que carezcan de ese “módulo” en su función de utilidad. En segundo lugar, si de lo que se trata es de maximizar utilidad, la propia dinámica de las experiencias vitales conduciría, especialmente en el caso de los GA’s beckerianos, *cæteris paribus* (y ya sea por aprendizaje individual o a través de procesos adaptativos), al propio oportunismo, la indiferencia ante el sufrimiento ajeno o la hipocresía. El problema de la ineficiente provisión de recursos destinados a la ayuda intraespecífica ha sido subrayado por diversos autores, entre ellos por Arrow (1972, p. 349): el maximizador de utilidad cuenta con incentivos altamente atractivos

15 Véanse al respecto Bruneau et al. (2012), y Gallese y Goldman (1998).

a no implicarse personalmente en la mejora de la utilidad ajena si pueden ser otros quienes lo hagan. Elster (2002, p. 497), por otra parte, critica la tesis del altruismo exclusivamente por conveniencia, o altruismo racional: si lo que motiva el comportamiento altruista es la búsqueda de una gratificación psicológica (*warm glow giving*) o eludir una emoción desagradable (como la culpa, o el estrés producto del contagio emocional), el agente racional bien puede invertir en recursos utilitarios alternativos o paliativos, tales como una imaginaria píldora anti remordimientos de conciencia, que aporten, igual que la estrategia de no mirar al indigente, un bienestar neto superior al que produce el altruismo.

## 6. Conclusiones

Parece que como ya explicara Keynes en el último párrafo de las Notas finales en su *Teoría general* (1936), los memes creados y propagados intelectualmente en el pasado continúan replicándose y evolucionando en vehículos de supervivencia que son las aparentemente nuevas teorías filosóficas y científicas. Los memes del egoísmo psicológico y del solipsismo epistemológico, asumidos y propagados en la tradición empirista británica, están presentes en la doctrina utilitaria de Bentham, y de ahí se replicaron -con algunas variaciones, especialmente en el caso del hijo- en la teoría psicológica de James y John Stuart Mill; en ambos apreciamos una continuidad esencial en los aspectos señalados. La teoría económica moderna ha recibido, igualmente con matices reconocibles si bien concordando en lo fundamental, esos memes evolucionados procedentes de la tradición utilitarista que inspiró su nacimiento y desarrollo.

El altruismo constituye un problema teórico en las ciencias del comportamiento que se basan en los supuestos del egoísmo psicológico y el alcance restringido de la intersubjetividad, que son dos caras de misma moneda. El egoísmo utilitario y las connotaciones asociadas al individualismo metodológico limitan la posibilidad de explicar toda la gama de comportamientos altruistas de la vida real. Las principales contradicciones y desafíos teóricos del altruismo explicado desde una visión utilitarista guardan relación con que en realidad no se está explicando un verdadero comportamiento altruista, sino un juego de palabras que esconde un altruismo práctico por conveniencia que en realidad se trata de egoísmo. En otro orden de cosas, la presencia de supuestos solipsistas en la teoría de la mente asumida plantea una incompatibilidad lógica no sólo con la práctica de la intersubjetividad, sino con la propia racionalidad asignativa de recursos a través del comportamiento altruista. La inverificabilidad de las representaciones mentales que constituyen la aproximación a toda experiencia ajena no se integra bien con la teoría evolutiva, dado que no parece razonable que la capacidad de realizar actos de altruismo haya evolucionado a ciegas de otras capacidades necesarias para guiar una conducta tendente a la eficiencia asignativa. El altruismo con resultados subjetivamente inverificables deriva bien en el exceso o defecto de transferencias altruistas de recursos, proliferación de conductas oportunistas tanto por el lado de la oferta como de la demanda de altruismo, así como en la evolución de la hipocresía conductual e indiferencia ante el sufrimiento ajeno. Estas conductas se hallan presentes en el etograma de la especie, y son estrategias de comportamiento adaptadas para la supervivencia. Sin embargo, también están presentes las conductas de un altruismo genuino, igualmente adaptativas, sin que podamos hablar con rigor de simplismos reduccionistas del tipo *homo aeconomicus* u *homo cooperatibus*.

En este sentido, la teoría del altruismo basada en la interdependencia utilitaria no sale bien parada de toda la problemática que emana del propio concepto de altruismo manejado, así como de los supuestos psicológicos y epistemológicos asumidos acerca del sujeto y su motivación. La psicología de inspiración milliana que porta en sus fundamentos, hemos comprobado, sigue limitando el alcance analítico y descriptivo del comportamiento humano, en este caso desde la teoría económica. El origen del problema estriba en la utilización de una teoría de la mente errónea, ampliamente superada hoy día en otras disciplinas académicas, y con limitaciones de alcance señaladas hace ya dos siglos y medio entre otros por Adam Smith y Adam Ferguson.

Ni los Mill ni Becker pudieron conocer los hallazgos recientes de la neurociencia cognitiva social, por lo que se precisa revisar el modelo económico general de explicación del altruismo en orden a solucionar sus deficiencias e incluir correctamente las experiencias del amor y la simpatía. La vía no explorada por los modelos inspirados en el utilitarismo, dígame que la frontera del yo se ve naturalmente amplificada en la experiencia simpatética y amorosa, permite asu-

mir la integración de las utilidades en un nexo cognoscitivo de acoplamiento y comunidad de sujetos. Este aspecto, sin embargo, plantea el problema de que no se concilia bien con el individualismo metodológico. La vía inexplorada fue uno de los ejes y tuvo un importante desarrollo en las obras de autores de la Escuela Escocesa de filosofía moral en el s. XVIII. Sin embargo, los memes de esta “vía escocesa” no han tenido tanta capacidad de supervivencia como los de la doctrina utilitarista... Hasta tiempos aún recientes.

## 7. Bibliografía

- Arendt, Hannah. 2002 [1978]. *La vida del espíritu*. Barcelona, Paidós.
- Arrow, Kenneth J. 1972. Gifts and exchanges, *Philosophy and Public Affairs*, vol. 1 (4), pp. 343-362.
- (1994) Methodological individualism and social knowledge, *American Economic Review*, vol. 84 (2), pp. 1-9.
- Ashraf, Nava, Camerer, Colin F., y Loewenstein, George. 2005, Adam Smith, behavioral economist, *The Journal of Economic Perspectives*, vol. 19 (3), pp. 131-145.
- Austin, John L. 1981 [1959]. Sentido y percepción. Madrid, Tecnos.
- Bartels, Andreas, y Zeki, Semir. 2004. The neural correlates of maternal and romantic love, *Neuroimage*, vol. 24 (3), pp. 1.155-1.166.
- Becker, Gary S. (1987). *Tratado sobre la familia*. Madrid, Alianza.
- 1997 [1974]. A theory of social interactions, *Journal of Political Economy*, vol. 82 (6), Traducción de Ramón Febrero y Pedro Schwartz (1997), pp. 177-205.
- 1997 [1992]. Habits, addictions and traditions, *Kyklos*, vol. 45 (3). Traducción de Ramón Febrero y Pedro Schwartz (1997), pp. 237-253.
- Blaug, Mark. 1985. *La metodología de la economía. O cómo explican los economistas*. Madrid, Alianza.
- Boulding, Kenneth E. 1971. Notes for a theory of philanthropy. En *Collected Papers*, Boulder, Colorado, Colorado Associated University Press, vol. II, pp. 235-249.
- Brown, Donald G. 1973. What is Mill's principle of utility? *Canadian Journal of Philosophy*, vol. 3 (1), pp. 1-12.
- Bruneau, Emile G., Pluta, Agnieszka, y Saxe, Rebecca. 2012. Distinct roles of the “shared pain” and “theory of mind” network in processing other's emotional suffering, *Neuropsychologia*, vol. 50 (2), pp. 219-231.
- Collard, David. 1975. Edgeworth's propositions on altruism, *The Economic Journal*, vol. 85 (338), junio, pp. 355-361.
- Copleston, Frederick. 2011. *Historia de la Filosofía*. Barcelona, Círculo de Lectores.
- Corsini, Raymond J. (ed.). 1999. *The Dictionary of Psychology*. Philadelphia & London, Brunner Mazel.
- Crisp, Roger. 1997. *Mill on Utilitarianism*. London, Routledge.
- De Juan, Oscar; y Monsalve, Fabio. 2006. Morally ruled behavior: The neglected contribution of Scholasticism, *European Journal of the History of Economic Thought*, vol. 13 (1), pp. 99-102.
- De Waal, Frans B. M. 2006. *Primates and philosophers: How morality evolved*. Macedo, Stephen, y Ober, Josiah (eds.); Princeton & New York, Princeton University Press.
- Edgeworth, Francis Y. 2000 [1881]. *Psicología matemática*. Madrid, Pirámide.
- Elster, Jon. 2000 [1979]. *Ulises y las sirenas. Estudios sobre racionalidad e irracionalidad*. México, Fondo de Cultura Económica.
- 2002. *Alquimias de la mente. La racionalidad y las emociones*. Barcelona, Paidós.
- Febrero, Ramón, y Schwartz, Pedro. 1997. *La esencia de Becker*. Barcelona Ariel.
- Ferguson, Adam. 1974 [1767]. *Un ensayo sobre la historia de la sociedad civil*. Madrid, Instituto de Estudios Políticos.
- Fontaine, Philippe. 2007. *From philanthropy to altruism: Incorporating unselfish behavior into economics, 1961-1975*, *History of Political Economy*, vol. 39 (1), pp. 1-46.
- Frank, Robert H. 1994. *Microeconomía y conducta*. Madrid, McGraw-Hill.

- Frantz, Roger. 2005. Two minds. *Intuition and analysis in the History of Economic Thought*, New York, Springer.
- Frey, Bruno S. & Stutzer, Alois. 2002. What can economists learn from happiness research? *Journal of Economic Literature*, vol. 40 (2), pp. 402-435.
- Gallese, Vittorio; y Goldman, Alvin. 1998. Mirror neurons and the simulation theory of mind-reading, *Trends in Cognitive Sciences*, vol. 2 (12), pp. 493-501.
- García Suárez, Alfonso. 1976. *La lógica de la experiencia. Wittgenstein y el problema del lenguaje privado*. Madrid, Tecnos.
- Ghiselin, Michael T. (1974). *The economy of nature and the evolution of sex*. University of California Press, Berkeley.
- Gregory, Richard L. (ed.). 2004. *The Oxford Companion to the Mind*. New York Oxford University Press.
- Harsanyi, John C. 1955. Cardinal welfare, individualistic ethics and interpersonal comparisons of utility, *Journal of Political Economy*, vol. 63 (4), pp. 309-321.
- Hayek, Friedrich A. 2004 [1952]. *El orden sensorial: Los fundamentos de la psicología teórica*. Madrid, Unión Editorial.
- Hein, Grit, y Singer, Tania. 2008. I feel how you feel but not always: The empathic brain and its modulation, *Current Opinion in Neurobiology*, vol. 18 (2), pp. 153-158.
- Iacoboni, Marco. 2009. *Las neuronas espejo*. Buenos Aires y Madrid, Katz.
- Jevons, William Stanley. 1998 [1871]. *La teoría de la política económica*. Madrid, Pirámide.
- Lamm, Claus, Batson, Daniel C., Decety, Jean. 2007. The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal, *Journal of Cognitive Neuroscience*, vol. 19 (1), pp. 42-58.
- Mill, James. 2001 [1829]. Analysis of the phenomena of the human mind. Reimpresión en dos vols. de la ed. de 1829, Bristol, Thoemmes Press.
- Mill, John Stuart. 1979 [1865]. An examination of Sir William Hamilton's philosophy and of the principal philosophical question discussed in his writings. En Robson, John M. (ed.), *Collected Works of John Stuart Mill*, London, Routledge & Kegan Paul, vol. 9.
- Nagel, Thomas. 2004 [1970]. *La posibilidad del altruismo*. México, Fondo de Cultura Económica.
- Robbins, Lionel C. 1938. Interpersonal comparisons of utility: A comment, *Economic Journal*, vol. 48 (192), pp. 635-641.
- Rutherford, Alasdair. 2010. Get by with a little help from my friends: a recent history of charitable organisations in economic theory, *The European Journal of the History of Economic Thought*, vol. 17 (4), pp. 1.031-1.046.
- Samuelson, Paul A. 1993. Altruism as a problem involving group versus individual selection in Economics and Biology, *American Economic Review*, vol. 83 (2), pp. 143-155.
- Saxe, Rebecca, y Kanwisher, Nancy. 2003. People thinking about thinking people. The role of the temporo-parietal junction in theory of mind, *Neuroimage*, vol. 19 (4), pp. 1.835-1.842.
- Searle, John R. 2000. *Razones para actuar. Una teoría del libre albedrío*. Barcelona, Círculo de Lectores.
- Sen, Amartya K. 1977. Rational fools: A critique of the behavioral foundations of economic theory, *Philosophy and Public Affairs*, vol. 6 (4), pp. 317-344.
- Shotland, R. Lance, y Stebbins, Charles A. 1983. *Emergency and cost as determinants of helping behavior and the slow accumulation of social psychological knowledge*, *Social Psychology Quarterly*, vol. 46 (1), pp. 36-46.
- Singer, Tania, Seymour, Ben, O'doherty, John, Kaube, Holger, Dolan, Raymond J., y Frith, Chris. 2004. Empathy for pain involves the affective but not sensory components of pain, *Science*, vol. 303, (5.661), pp. 1.157-1.162.
- Smith, Adam. 1997 [1759]. *La teoría de los sentimientos morales*. Madrid, Alianza.
- Sobel, Joel. 2005. Interdependent preferences and reciprocity, *Journal of Economic Literature*, vol. 43 (2), pp. 392-436.
- Sober, Elliott. 1996. *Filosofía de la biología*. Madrid, Alianza.
- Sober, Elliott, y Wilson, David Sloan. 2000. *El comportamiento altruista. Evolución y psicología*. Madrid, Siglo Veintiuno de España.

- Spikins, Penny A., Rutherford, Holly E., y Needham, Andy P. 2010. From hominity to humanity: Compassion from the earliest archaic to modern humans, *Time and Mind*, vol. 3 (3), pp. 303-325.
- Thaler, Richard H. 1992. *The winner's curse: Paradoxes and anomalies of economic life*. Princeton (NJ), Princeton University Press.
- Trincado, Estrella. 2009. *Crítica del utilitarismo. Utilidad frente a realidad presente*. Madrid, Maia.
- Tyler, Tom R., y Blader, Steven L. 2000. *Cooperation in groups. Procedural justice, social identity, and behavioral engagement*. New York, Psychology Press.
- Vignemont, Frederique y Singer, Tania. 2006. The empathic brain: How, when and why? *Trends in Cognitive Sciences*, vol. 10 (10). pp. 435-441.
- Watson, John. 1966. *Teorías del placer*. Buenos Aires. Paidós.
- Wilson, Edward O. 1980[1975] *Sociobiología: La nueva síntesis*. Barcelona, Omega.
- Wittgenstein, Ludwig J. 1997 [1921]. *Tractatus logico-philosophicus*. Madrid, Alianza.