

EL LIMITADO FRACASO DEL *HOMO* *OECONOMICUS*

Félix Ovejero

Universidad de Barcelona

Sumario: 1. Introducción. 2. Ventajas del *homo oeconomicus*. 3. El *homo oeconomicus* como racionalidad. Tres interpretaciones. A. Estrategia heurística. B. Competencia cognitiva. C. Teoría empírica. 4. Los problemas del *homo oeconomicus*. A. El egoísmo. B. Sesgos, heurísticos e incoherencias. C. Simplicidad. D. Incapacidad social. 5. Conclusión: una valoración matizada.

1. INTRODUCCIÓN

“Verde es el árbol de la vida, gris el de la inteligencia”. La fatigada sentencia de Goethe se puede entender de muchas maneras. Entre ellas, como el resignado reconocimiento de que en ciencia nos enfrentamos a un complicado equilibrio entre la conveniente parsimonia de nuestras teorías y las realidades que queremos entender, siempre más complejas. La ciencia apuesta por la simplicidad. Un buen ejemplo de ello es el *homo oeconomicus* (HE, en lo que sigue), el patrón antropológico más popular de la teoría social, racional y egoísta. La idea es austera y poderosa: con poco explicaría mucho. También parece realista, compatible con nuestra experiencia cotidiana: no somos héroes morales, más bien al contrario. Viene, además, avalada por la teoría social más reputada, la ciencia económica. Está en la trastienda de sus explicaciones, tanto de las microeconómicas, cuando da cuenta de los precios como resultado de las interacciones entre consumidores y productores, unos maximizando su utilidad y otros su beneficio, como de las macroeconómicas, como sucede cuando relaciona una subida de los tipos de interés con una caída de la inversión y de la demanda agregada (mediante el supuesto de que los agentes optan por man-

tener el dinero en el banco y, además, dado su precio, no solicitan créditos).

Pero hay más. Con ese guión, se dice, podemos entender todo lo que tenga ver con los humanos en sociedad: las personas no se casan por amor ciego, sino porque, llegado un punto, el coste marginal de seguir explorando comienza a ser más alto que el beneficio previsible; los criminales no son malos tipos o producto de familias desordenadas, sino refinados calculadores que comparan los costes y los beneficios de su actividad; los políticos no son ciudadanos virtuosos, sino maximizadores de votos; los emigrantes, cuando envían dinero a sus países de origen, no lo hacen por amor a sus familias sino para evitar la llegada de compatriotas y la consiguiente disminución de sus salarios; las personas provistas de talentos o habilidades no están dispuestas a ejercerlos si no obtienen una retribución superior a los demás, de modo que incluso la sociedad más justa tendrá que propiciar la desigualdad; los ricos contribuyen a la beneficencia para comprarse un sentimiento de superioridad moral; los niños mejoran su rendimiento escolar si les pagamos por leer libros.

Este programa, si se cumple, dejaría en nada a la mecánica de Newton que, total, se limitaba a explicar las mareas, la trayectoria de

los planetas, la caída de los cuerpos y cuatro cosas más. El HE daría cuenta de nuestro comportamiento de la cuna a la tumba y de la cama a la banca. En las páginas que sigue haremos una valoración de tan optimista consideración. En primer lugar se abordarán las razones de su éxito. Se verá, también, que, en realidad, esa potencia explicativa tiene que ver con una cierta ambigüedad del personaje que se desenvuelve entre una simple estrategia explicativa, una competencia cognitiva y una genuina teoría empírica. Cuando oficia como estrategia explicativa resulta trivialmente compatible con diversos comportamientos, pero, en rigor, no explica ninguno. Cuando se interpreta como una competencia cognitiva, comparable a la visión o la memoria, nos enfrentamos a la posible existencia de un módulo mental que más que explicar necesita ser explicado. Cuando se entiende con una genuina teoría pierde buena parte de su vigor explicativo. Se examinarán algunos de los resultados que ponen en entredicho el potencial del HE. Con todo, se argumentará en favor de una interpretación prudente del alcance de esos resultados o, lo que es lo mismo, en contra de una descalificación absoluta del personaje. Quizá no sea el protagonista de la obra, pero sí un secundario importante.

2. VENTAJAS DEL HOMO OECONOMICUS

El triunfo del HE se produce en un doble plano: en el mundo y en las teorías sobre el mundo. Es posible que sus virtudes teóricas y epistémicas expliquen su vocación de monopolio entre los distintos modelos antropológicos que han nutrido la teoría social, pero lo que es innegable es que ese triunfo en el terreno de las ideas fue precedido de un triunfo en el plano material, en la historia. Entre las diversas motivaciones humanas, las nacientes sociedades capitalistas alentaban la búsqueda del interés. Se nutrían del HE y propiciaban su

proliferación¹. En ese sentido, se puede hablar de un desplazamiento del combustible social desde las pasiones hacia los intereses². Un desplazamiento ventajoso en más de un sentido.

Los moderados burgueses buscan sus beneficios y no se entregan al activismo político, las emociones o las religiones. La fórmula “el dinero es cobarde y previsible” condensa algunas de esas ventajas. Por una parte, el *doux commerce* “suaviza las costumbres bárbaras”, favorece la tolerancia mediante el trato con otros, alienta virtudes como la moderación, la honradez, la honestidad, la frugalidad y la laboriosidad³. Además, en el mercado, cuando cada uno procura por sí mismo de un modo calculado y, en competencia, se ve obligado a atender las demandas de los consumidores, no se embarca en otras empresas que las que se traducen, sin que nadie lo pretenda, en la prosperidad para todos: la mano invisible que popularizará Adam Smith⁴ y que anticipaba Vico: “las pasiones de los hombres ocupados por entero a las búsqueda de su utilidad privada se transforman en un orden civil que permite a los hombres vivir en sociedad humana”⁵. Sobre el sustrato de ese orden civil se pueden levantar las instituciones. El egoísmo de las gentes asegura un orden y permite orientarse a los gobernantes. James Steuart, el padre de la economía política —al menos el acuñador de la fórmula, nueve años antes de la publicación de *La riqueza de las Naciones*— lo resumía con eficacia, cuando, después de recordarnos que “si todos los días ocurriesen milagros, no habría leyes”, acotaba: “si todos actuaran en favor del interés público y se olvidaran de sí mismos, el estadista se sentiría desconcertado”⁶.

Ese será el orden del mundo y sobre su posibilidad se hará posible la vida compartida y la naciente teoría social. Con algún detalle, el HE aseguraría:

1. Orden ontológico. Para una parte del pensamiento social la tesis de que cuando cada uno va a la suya se produce un orden social será la piedra basal en la que fundamentar el conocimiento. La mano invisible, en este caso,

se entenderá como una suerte de tesis ontológica o metafísica equiparable al principio determinista: sin la regularidad del mundo, sin asumir que las mismas causas producen los mismos efectos, no habría posibilidad de experimentación repetida ni de conocimiento. En sentido parecido, de la interacción de los HE emergería un orden regular, a la vez que espontáneo, sin necesidad de objetivos comunes ni planificados entre los agentes⁷. Esa inexorabilidad se nutre, llamativamente, de las elecciones de cada persona, cuando opta por ir a la “suya”. Una estrategia, la de procurar el interés propio, que, además de ser la dominante, la que asegura el mejor resultado, hagan lo que hagan los demás, es, a la postre, la única posible: en un escenario competitivo, comportarse con generosidad, cuando los otros atienden a su interés, es el camino más seguro a la desaparición. Es posible que todos quieran optar por otro proceder, pero saben que si lo hacen, serán los primeros en caer. Así que mejor anticiparse. Basta con que piensen que los demás intentarán buscar su particular beneficio —sea verdad o no— para que el mecanismo opere. El individuo “elige” hacer lo que no le queda más remedio que hacer, qué es, además, lo único que puede hacer: buscar su beneficio. El orden de todos sostenido en la libre elección de cada cual.

2. Orden moral. El orden de mundo no implica el orden moral. El buen conocimiento físico, sostenido en la regularidad de mundo, nos permite anticipar la muerte del Sol, que será, sin duda, un equilibrio: el núcleo solar se transformará en una enana, que, mucho más tarde, al enfriarse, derivará en una enana negra y entonces se acaba la historia, el equilibrio final. También resultaron equilibrios las catástrofes ambientales de distintas sociedades (los isleños de Pascua, los vikingos en Groenlandia, los habitantes de las islas Pitcairn, Henderson y Mangareva) que tan magníficamente nos describe Jared Diamond⁸. Para muchos el HE produce otro tipo de equilibrios: equilibrios buenos, interesantes según algún criterio normativo. El mercado es ejemplo paradigmá-

tico, pero no el único. Sucede en muchas otras ocasiones, por ejemplo, cuando al caminar por un bosque optamos por la senda que otros han transitado antes que nosotros: no queremos hacer bien, ni mal, a nadie, pero con ello contribuimos a que otros, que vendrán después, puedan caminar con más facilidad. Además, si queremos llegar lo más pronto posible, elegiremos ese camino. Cada uno con su acción, sin quererlo y sin coordinarse con los demás, facilita la vida a los demás⁹. En el mercado, la competencia, me obliga a ir a la mía. Lo acabamos de ver: los agentes eligen lo que no pueden dejar de elegir. Pero es que, además, eso que no les queda más remedio que elegir, es lo que realmente quieren elegir, porque está en su naturaleza, el *Homo homini lupus* de Plauto. Todavía más, es lo que deben elegir, añadirá la metáfora de la mano invisible: lo que asegura el beneficio social. Lo que quiero, lo que puedo y lo que debo en una sola operación.

3. Coordinación en la interacción social. Saber que los humanos buscan exclusivamente su beneficio ayuda a anticipar su comportamiento. Algo importante en dos planos, al menos. Primero, en las acciones entre las personas: yo sé que tú, que conduces en dirección contraria a la mía, no tienes intención de cambiar de carril “por la cuenta que te trae” y tú te mantendrás en tu carril porque conjeturas lo mismo acerca de mí. Segundo, en las instituciones, en los términos que sintetizaba Steuart: si aumentan las penas, disminuyen los delitos; si bajan los tipos de interés, aumenta la inversión; si subo las pensiones, me votarán los jubilados. Mientras una sociedad asentada en las pasiones, las ambiciones o las ideologías resulta incierta y, con una alta probabilidad, requiere costosas intervenciones, exhortaciones, vigilancias o represiones, el egoísmo abarata la ingeniería social. Lo dicho: el dinero es cobarde y previsible, siempre que sea racional y egoísta.

4. Predictibilidad teórica. el HE tiene una única motivación, su interés particular, y solo quiere un resultado, el mejor. No se dispersa ni se desorienta con emociones, ideologías o religiones y, además, entre todas las opciones

compatibles con sus intereses, solo juzga una como buena, la óptima. Esa es la perspectiva de la primera persona descrita en el punto anterior, la que opera en las relaciones entre los agentes y sobre ésta se sostiene la perspectiva del estudioso, de la tercera persona. Pero, además, si la motivación es el máximo interés, el análisis puede mejorar. Una unidimensionalidad apunta a la posibilidad de establecer una función métrica que asocie números a las conductas en una escala de utilidad. Desde el punto de vista de la teoría social eso quiere decir que podemos afinar en nuestras predicciones: mientras hay muchos “suficientes”, solo hay un punto óptimo, donde se maximizan los beneficios. Unos móviles de conducta que no cambian con sociedades, religiones o culturas suponen una apuesta por teorías parsimoniosas, con pocos supuestos, y alto contenido empírico.

3. EL HOMO OECONOMICUS COMO RACIONALIDAD. TRES INTERPRETACIONES

En principio, la idea de HE, egoísta y racional, gana en precisión cuando nos movemos en el último de los terrenos mencionados: la maximización del beneficio. La mejora parece todavía mayor cuando la idea de beneficio se sustituye por la de utilidad y, aún más, cuando la utilidad se entiende, como hace la moderna economía, como una función asociada a un orden de preferencias. Eso, al menos, es lo que parece. Porque por ese camino también aumenta el riesgo de vaciedad, de que el HE acabe por resultar compatible casi con cualquier comportamiento. Como escribía Amartya Sen: “si la persona es coherente, se trate de un obtuso egoísta o de un admirable egoísta o de un defensor consciente de sus propios intereses de clase, su comportamiento aparecerá orientado a la maximización de su utilidad en este mundo encantado de las definiciones”¹⁰. En realidad, con esa recalificación, los com-

ponentes originales del HE, egoísmo y racionalidad, si no se añaden más consideraciones, acaban por resultar inseparables y hasta por confundirse: cualquier cosa que el sujeto desee, por definición, maximizará su utilidad y se podrá entender como su interés. Lo que encabeza el orden de preferencias, coherente, por definición, se identificará con lo que interesa¹¹. Vista así, la idea de “orden de preferencias” no hace más que reunir las dos ideas: racionalidad (coherencia, esencialmente transitividad) y egoísmo (en tanto, se persiga lo que se persiga se entenderá como aquello que maximiza la utilidad). Al final, la idea de HE se reduce a la de racionalidad. Normalmente, para evitar deslizarse por esa pendiente, más pronto o más tarde, la teoría económica, en la medida en la que aspira a ser una ciencia empírica, a explicar, acaba por añadir suposiciones adicionales, cualificaciones para dotar a la idea de contenido informativo.

Desde otro punto de vista, la falta de contenido informativo, esa conversión en simple racionalidad se puede entender como indicación de que estamos ante diversas interpretaciones o funciones de la idea de HE. La ambigüedad afecta al estatuto del esquema R: Para todo X, si X desea D y X cree que, teniendo en cuenta todas las circunstancias, la acción A es la más eficiente para obtener el objetivo D, entonces X hace A¹². Para unos el esquema R se entenderá en un sentido parecido al principio causal, como una simple estrategia explicativa o heurística eficaz para ordenar una parte de nuestras experiencias. Para otros R es el programa que rige un módulo epistémico con el que los humanos tratamos con la realidad, incluida las relaciones con los demás: una facultad cognitiva como la visión o la memoria. El agente interpreta las acciones de los demás atribuyéndoles motivos, razones, unas creencias y unos deseos. En un tercer sentido R se entenderá como una genuina teoría empírica susceptible de control experimental, como pueden serlo las leyes de Mendel. Apostar por una de esas interpretaciones no conlleva excluir las otras. En realidad, la teoría económica, cuando dota

de contenido al HE, perfila la tercera idea y estipula una teoría de la racionalidad muy específica susceptible de control empírico, asume que esa teoría describe lo que realmente hacen los agentes (segunda idea) y ante cualquier acción A opta por abordarla (primera idea) como un problema de intereses e incentivos: ¿qué objetivo (interés) D perseguían los que hacían A?. Veamos con algún detalle las tres interpretaciones:

A. ESTRATEGIA HEURÍSTICA

Al explicar por qué las tribus del actual Estado de Montana cazaban a los bisontes en otoño e invierno y por qué preferían las hembras a los machos, el paleontólogo opta por asumir la racionalidad de los agentes. Sobre ese supuesto, levanta su conjetura: “era una respuesta antes de la llegada del crudo invierno y escogían hembras porque separadas con sus crías de los machos suponían una presa fácil”. Si el registro fósil mostrara otra cosa, por ejemplo, que los sacrificios se daban en primavera y que recaían en machos —como sucedía en Nuevo Méjico—, el investigador no dirá que “estaban locos”. Sin abandonar el supuesto racionalidad, buscará otras razones, con R como guión incuestionable. Conjeturará, por ejemplo, que esa elección respondía a que un régimen demasiado rico en proteínas supone un consumo metabólico muy importante, que, aunque para organismos bien nutridos no plantea problemas, sí los tiene para unos cazadores recolectores que siempre podrían compensarlo mediante la absorción de grasas o hidratos de carbono¹³.

En este caso, R es una esquema que nos “permite ordenar los datos, explicar interrelaciones y realizar preguntas a la Naturaleza”¹⁴. Opera como una red analítica útil para capturar ciertas piezas de la realidad mediante ciertas preguntas. En otros contextos podremos utilizar otras redes. Tres son las más eficaces. La primera, la red causal o enfoque físico que lleva a preguntarnos, por ejemplo, porque el agua burbujea al hervir. La segunda, la red funcional

o enfoque adaptacionista (o del diseño), particularmente fecunda al enfrentarse a los procesos evolutivos, apuesta por preguntarse para qué sirve cierto órgano, por la función que cumple: el problema adaptativo que resuelve y para el que parece estar diseñado. La tercera, la de R, conjetura los intereses que están detrás de las acciones e interpreta el comportamiento de un ente (persona, animal, artefacto, cualquier cosa) como si fuera un agente racional que “elige” sus “acciones” “teniendo en cuenta” sus “creencias” y sus “deseos”¹⁵.

En este caso no estamos ante explicaciones sino ante esquemas de explicaciones¹⁶. Si no somos capaces de explicar una enfermedad, de dar con sus causas, no decimos que no tiene causa sino que no la encontramos. La investigación avanzará asumiendo que hay una causa y las preguntas se harán bajo el horizonte incuestionable de la red causal. Dos médicos discreparán acerca de la causa de la enfermedad, pero no del supuesto de qué tiene una causa. Lo mismo sucede con esta idea de racionalidad. Discreparemos acerca de las razones de los agentes, de sus creencias o de sus objetivos, pero sin abandonar el guión R. No dudamos de la red “creencias y deseos”, aunque podemos dudar de si cierta creencia o cierto deseo son los que realmente explican el comportamiento del agente, de las conjeturas empíricas. Si acaso, en determinado ámbito, abandonaremos el guión porque resulta poco fecundo, porque no nos ayuda a hacer preguntas, pero no podemos decir que “es falso”, a lo sumo, inútil. Resulta difícil en este contexto hablar de comportamientos irracionales¹⁷.

En teoría social es casi imposible prescindir de esa estrategia¹⁸, aunque también puede ayudarnos en otros ámbitos: un perro que desea aplacar su sed se dirige a un barreño de agua y bebe; una rata que elige el trozo de queso de mayor tamaño; las hormigas —que carecen de una visión global del territorio— cuando (siempre) escogen el camino más corto. También un calefactor provisto de un termostato: detecta (“percibe”, “cree que hay”) un desajuste entre la temperatura “deseada” y la real y se

pone en marcha para “conseguir” su objetivo, la temperatura ideal. En principio, no existirían límites al uso de esas estrategias interpretativas ni por tanto ámbitos de aplicación exclusiva; sencillamente sucede que en unos casos una resulta más eficaz que la otra. A la hora de explicar las acciones lanzamos una red analítica que apela a las “razones” de los actores, a sus creencias y sus objetivos. Hay un objetivo y se actúa para conseguirlo a partir de cierta información, esto es, de creencias sobre cómo son las cosas. No tenemos por qué asumir que “realmente es así”, que el termostato “cree” o “quiere”. R nos proporcionaría una simple pauta. Si funciona, adelante.

B. COMPETENCIA COGNITIVA

El conductor de un coche trata de llegar en el menor tiempo posible (o por la carretera más hermosa o con el menor consumo o por la ruta más exigente), dadas ciertas limitaciones (un cierto nivel de riesgo, de consumo, de habilidad, etc.) y de cierta información (sobre las rutas, el estado de su auto, el tiempo, el funcionamiento de los semáforos). En ese sentido es racional. Tiene objetivos y creencias sobre el mundo. Pero también lo es en otro sentido. Tiene creencias sobre el comportamiento de los demás. Supone que el peatón respetará el semáforo, que otro conductor, cuando encienda en intermitente, cambiará de carril; que el camionero no acelerará cuando intenté adelantarle. Asumirá que son racionales y ellos harán lo mismo con él. Tienen creencias sobre la creencias de los demás.

Las personas operamos como “psicólogos naturales”. En nuestros tratos, nos atribuimos intenciones, a las que damos respuestas, nos asumimos como individuos y, en ese proceder, adquirimos conciencia reflexiva. Tenemos, se dice, una “teoría” de la mente en la que ocupa un lugar central el supuesto de mutua racionalidad. Nos permite explicar, predecir y hacer inteligible el comportamiento de los demás sobre la base de sus creencias y deseos, del supuesto, por mejor decir, de que tienen

creencias y deseos, del supuesto de su “mente”. Se trataría, según algunos, de una “teoría implícita”.

En rigor, en tanto una teoría está compuesta por enunciados, la idea de teoría “implícita” carece de sentido. Tiene que ser enunciada, y, en ese sentido, es mejor hablar de un talento, de una competencia cognitiva con una base biológica, material, como la memoria o la visión. Se trataría de circuitos mentales evolucionados especializados en resolver problemas de coordinación social, sin necesidad de consciencia, instrucción formal y aplicados sin conocimiento de la lógica subyacente. Serían talentos “obvios” que nos permiten el razonamiento social en el mismo sentido en el que las inferencias necesarias para la ecolocalización son obvias para los murciélagos¹⁹. Tales talentos se han asentado en nuestro cableado mental por su funcionalidad adaptativa: en el pleistoceno, en grupos con alto grado de estabilidad, con individuos con largos procesos de maduración, con periodos improductivos y convivencia de miembros de distintas generaciones, resultaba importante realizar inferencias sobre sus actitudes, motivos y anticipar la futura conducta de los demás²⁰.

Esa capacidad opera en todas nuestras prácticas sociales²¹. Se da en la competencia, en un mercado, al anticipar los comportamientos de los rivales y de los clientes, y también en la colaboración, como sucede en la comunicación, que se basa en la expectativa de que el otro transmite con veracidad, brevedad, claridad y orden. El oyente asume que la información es relevante, no le es conocida y está suficientemente relacionada con lo que conoce como para realizar esa inferencia que es la comunicación: inferencia que le permite interpretar las intenciones del otro, eliminar ambigüedades, ordenar informaciones parciales y completar fragmentos (pronombres, alusiones, etc.)²². La más mínima comunicación opera sobre la existencia de esa capacidad de mutua atribución de estados mentales. Entiendo que quieres que te aproxime el salero cuando me señalas el salero o a que vas a desplazar el coche

cuando enciendes el intermitente y, a la vez, tu señalas el salero o enciendes el intermitente porque asumes que yo te atribuyo intenciones. Tu quieres que yo crea o sepa que tu quieres el salero o cambiar de trayectoria.

El soporte “yo pienso que tu piensas que yo pienso que...” es un cimiento de las relaciones humanas. Se da en la trama de la normas y en su violación, en la confianza y también en la mentira, porque cuando me engañas tu quieres que yo piense que tu piensas A cuando, en realidad, piensas B²³. Se da también en las emociones, mejor dicho en el reconocimiento de las emociones, que presume una capacidad de inferir los estados mentales de las otras personas, de tener pensamientos sobre pensamientos (o emociones), del tipo “ese otro siente (cree, quiere) x o y”, “ese otro siente (cree, quiere) que yo sienta (crea, quiera) x o y” o “ese otro siente (cree, quiere) que yo sienta (crea, quiera) que él siente (cree, quiere) x o y”, etc²⁴.

Esa capacidad resulta decisiva en escenarios de reciprocidad en tanto permite realizar inferencias complejas sobre quienes pueden socavarla. Hace posible la mentira pero también detectar a los mentirosos y a los *free riders*, reconocer el fraude²⁵. Facilita el desarrollo de la cooperación y, a la vez, la amenaza y la venganza. El juego emocional depende “de la naturaleza de la máquina interpretativa de la especie”²⁶. De la especie porque, para ser eficaz, la capacidad tiene que ser común a todos. Si el otro no es capaz de categorizar nuestras acciones de represalia por su conducta incorrecta, el castigo resulta inútil. Es lo que sucede, seguramente, con los autistas, que, aun si inteligentes, resultan particularmente indefensos socialmente. Carecen de la capacidad de representarse las representaciones mentales de los otros y, en consecuencia, se muestran incapacitados para comprender el significado normativo de los acontecimientos. Si los humanos a veces exageramos atribuyendo intenciones a la naturaleza (ese “los dioses nos castigan con una plaga por nuestros pecados”), los autistas no llegan nunca. Ellos tratan a los demás como

nosotros a la naturaleza (salvo en las sociedades animistas).

C. TEORÍA EMPÍRICA

Lo que encontramos en los libros de economía, y se extiende al conjunto de la teoría social, es otra cosa²⁷. Se trata de una genuina teoría, formulada en una serie de enunciados (explícitos) y susceptible de control empírico²⁸. La idea más precisa del HE. Eso no quiere decir que sea unívoca. Las formulaciones son diferentes, con distintos grados de exigencia, según los sistemas reales que se quieren explicar. E incluso hay diferencias en la presentación del núcleo de la teorías. Así, por ejemplo, en uno de los textos pioneros en la extensión de las teorías económicas a otros dominios de la realidad social, la hipótesis del HE se asocia a la suposición de que “Todos los comportamientos humanos se pueden entender como envolviendo a participantes que (1) maximizan su utilidad (2) a partir de un conjunto de preferencias estables y (3) recogen una óptima cantidad de información y de otros *inputs* en una variedad de mercados”²⁹. Por su parte, en uno de los pocos manuales de microeconomía sensibles a los resultados de la psicología económica, se da una versión ligeramente distinta, según la cual, “ser racional significa tomar decisiones de acuerdo con el criterio coste-beneficio, es decir, realizar una actividad si y solo si los beneficios son superiores a los costes. Esta definición de la racionalidad puede ser objeto de dos refinamientos importantes. Uno es el criterio de la racionalidad basado en el egoísmo, según el cual las personas racionales conceden un gran peso únicamente a los costes y beneficios que les afectan directamente a ellas (.) La otra definición es el llamado criterio de la racionalidad basado en el objetivo inmediato. Su único requisito es que las personas actúen eficientemente de acuerdo con las aspiraciones y objetivos que tengan en cada momento”³⁰.

La diversidad de interpretaciones está vinculada al énfasis en distintos aspectos de la ra-

cionalidad económica: como egoísmo y como consistencia interna, entendida esta última como aquella elección en la que “las diferentes cosas escogidas de diferentes subconjuntos se corresponden unas a otras en un modo plausible y coherente”³¹. A esas interpretaciones, en otras ocasiones Amartya Sen añade una tercera que, en rigor, es una prolongación de la racionalidad como consistencia: la racionalidad como comportamiento maximizador, que en una formulación intuitiva se corresponde con la hipótesis de que los individuos, dadas sus restricciones (de información, de recursos y técnicas), escogen la acción que mejor les asegura la obtención de sus objetivos, y en otra más precisa, se asocia con la teoría de la maximización de la utilidad esperada, según la cual, los individuos, que tienen preferencias estables y consistentes, escogen la alternativa que maximiza el valor esperado de su función de utilidad³².

Por detrás de las diferentes presentaciones se encuentran un conjunto de axiomas compartidos con pequeñas variaciones, un núcleo mínimo de supuestos que se van modificando o ampliando en función de los requerimientos explicativos: según las creencias (bayesianas, capturando una cantidad “óptima” información), los tipos de escenarios (certidumbre, incertidumbre, riesgo) y las exigencias formales (que permiten introducir unos tipos de representación con funciones continuas o unas métricas más o menos informativas (ordinales o cardinales)) y materiales sobre las preferencias (completitud, transitividad, independencia de las alternativas irrelevantes)³³. A partir de tales axiomas, sería posible, en principio, establecer las funciones de utilidad de los agentes. Adicionalmente, y aunque esto no siempre se hace explícito, se supone que los individuos actúan según tales axiomas, esto es, no solo tienen unas preferencias y creencias con ciertas características, sino que, además, los agentes actúan a partir de ellas³⁴ o, más exactamente, sus acciones se explican a partir de ellas³⁵. Finalmente, aunque la función de utilidad nada nos dice acerca del contenido de las preferen-

cias, de qué es lo que quieren las personas, de un modo u otro, *de facto*, siempre se asume que actúan procurando su propio interés, que son egoístas, en el sentido más común de la palabra egoísmo, como búsqueda de beneficios materiales³⁶. En breve, se hacen suposiciones respecto a las creencias, las preferencias, el egoísmo y las relaciones entre creencias y preferencias y acciones³⁷.

La descrita es, en pocas líneas, la formulación más consumada del HE, como una genuina teoría empírica. Estamos ya ante una precisa idea de persona. Y por eso mismo cabe preguntar si se ajusta con lo que sabemos sobre el comportamiento de los individuos reales. Después de todo, su aval fundamental no es otro que el realismo. Realismo en el sentido más o menos técnico de la filosofía de la ciencia, como descripción ajustada de las características básicas del comportamiento humano, y también realismo en el sentido más común, casi vulgar, de “no pedir peras al olmo”, de desconfianza respecto a la bondad humana, ese que queda recogido en refranes y proverbios del tipo “con dinero baila el perro”, “frailes y monjas, del dinero esponjas”, “más vale dinero en mano, que ilusiones de verano” o “no hay tan buen compañero como el dinero”.

4. LOS PROBLEMAS DEL HOMO OECONOMICUS

Ciertamente, no hace falta alejarse mucho de nuestras experiencias comunes para poner en duda el realismo del HE, su racionalidad y su egoísmo. Los ejemplos abundan. No somos simples egoístas y hasta parece que ni somos coherentes. Incluso hay razones para dudar de que seamos mínimamente racionales. Nos resulta más difícil desprendernos de un billete que de una moneda del mismo valor. No respondemos igual ante un producto con la etiqueta “95 % libre de materia grasa” que ante otro con la etiqueta “5 % de grasa”. En no pocas ocasiones, a pesar de que preferimos X a Y,

cuando, de pronto, nos añaden la posibilidad de elegir Z, que no nos interesa, cambiamos y pasamos a preferir Y a X. Si nos dicen que nuestro vecino es de tez pálida, bajito y come mucho arroz, creemos que es más probable que sea catedrático de chino que conductor de autobús, ignorando que, dado el número de los dedicados a tales actividades, lo más probable es exactamente lo contrario.

Y esos comportamientos no son extravagancias o rarezas. De hecho, son lo bastante regulares como para que las empresas o los gobiernos los tomen como punto de partida para sus políticas. Los tienen en cuenta los comerciantes cuando presentan sus productos o sus precios, los políticos cuando contemplan la posibilidad de contener la inflación sustituyendo monedas por billetes, los dueños de *Starbucks* cuando nos ofrecen como señuelo —ese Z descartado— el recipiente supergrande para que elijamos el “mediano” antes que el pequeño, que quizá era nuestra primera opción, y los guionistas de series televisivas cuando quieren orientar nuestras sospechas³⁸.

Aunque las dificultades afectan a cada uno de las dimensiones antes mencionadas (creencias, preferencias, egoísmo y relaciones entre creencias y preferencias y acciones), aquí utilizaré el guión clásico e inventariaré los problemas afectan a los dos aspectos típicos del HE, el egoísmo y la racionalidad³⁹. Añadiré, además, una mención a su simplicidad psicológica que trata de capturar resultados que se quedan a medio camino entre el egoísmo y la racionalidad y otra que apunta a su incompatibilidad con la vida social. Esta última, la más provisional, es la dificultad más crítica: si el HE resulta incompatible con la vida social, la simple existencia de ésta sería un argumento definitivo en contra de aquél.

A. EL EGOÍSMO

Cuando hemos de invertir nuestros ahorros, elegir un trabajo o comprar un piso nos comportamos como el HE. Atendemos a

los tipos de interés, los salarios y los precios. Pero también tenemos en cuenta otras consideraciones: la localización o el trato que nos dispensan en la oficina bancaria, la naturaleza y el modo de realizar la actividad laboral, la proximidad a ciertos servicios o a las casas de nuestros amigos. Por supuesto, los economistas nos dirán que, en todos esos casos, cabe “endogeneizar” esas otras variables, reducirlas a interés. Las funciones de utilidad condensarían en un valor único las distintas dimensiones: el tipo de interés, la localización y el trato de la oficina se ponderarían e integrarían en un valor único y, una vez se tiene todo en cuenta, se pueden ordenar las opciones y decidir. Y, en cierta medida, tendrán razón los economistas, al cabo, con frecuencia, hacemos valoraciones globales. Por ejemplo, al elegir un restaurante consideramos el precio, la comida, los vinos, el trato, el ambiente y —eso sí, sin afinar ni ordenar explícitamente— optamos entre ellos. Pero no es menos cierto que, en otros casos, esa traducción resulta mucho más difícil, al menos si no queremos acabar en vaciedades tautológicas a cuenta de la utilidad, como las de aquel personaje de Molière que “explicaba” los efectos narcotizantes del opio apelando a una virtud dormitiva, cuya existencia se probaba por el hecho mismo de producir sueño. En todo caso, los experimentos de la psicología económica o economía experimental nos muestran que en nos parecemos bien poco al egoísta HE.

El juego del *ultimátum* es, seguramente, el resultado más conocido⁴⁰. En ese juego hay que dividir una cantidad de dinero entre dos participantes. El primero hace una oferta que el segundo acepta o rechaza. Si es aceptada, ambos se llevan la parte acordada; pero si el segundo la rechaza, los dos se quedan sin nada. Un sujeto egoísta, que asuma que el otro también lo es, propondría una miseria suponiendo que el segundo pensará que “mejor algo que nada”. Pues bien, los experimentos muestran que los primeros jugadores hacen propuestas no muy alejadas de la igualdad y que los segundos prefieren quedarse con nada a divisiones radicalmente desiguales. Otro experimento

se refiere a un caso de bienes públicos, bienes de cuyo consumo no hay modo de excluir a nadie: los individuos disfrutaban de la limpieza del aire o de la defensa nacional tanto si contribuyen a financiarlas como si no. Es sabido que la producción de tales bienes se enfrenta a un problema que complica su obtención, al menos si nos encontramos entre HE: cada individuo prefiere que los demás colaboren en su obtención mientras él se abstiene de todo esfuerzo y, por tanto, el bien público no se consigue. En el experimento, realizado a través de varias rondas sucesivas, a 10 individuos se les dice que recibirán 1 dólar cada uno en su cuenta por participar (haga lo que haga) en cada ronda. Por cada dólar que cualquiera de ellos aporte a la comunidad, cada uno recibirá $\frac{1}{2}$ dólar en su cuenta particular. El proceso se repite 10 veces y al final cada uno se lleva lo que tenga en su cuenta. Si los individuos cooperan y cada uno de ellos pone 1 dólar en la cuenta común, ésta tendrá 10 y en la cuenta de cada uno de ellos habrá, en cada turno, 5, que serán 50 al final. Si un sujeto es egoísta, guardará su dinero y recibirá de acuerdo con lo que aporten los demás. Si todos los demás son cooperativos, el egoísta se quedará con 55 y los otros con 45. Si todos son egoístas, cada uno se quedará con 10, uno por cada partida. Cuando el experimento se lleva a cabo, los resultados muestran que muy pocos adoptan la estrategia egoísta, que, por el contrario, prima la cooperación⁴¹.

En realidad, no hace falta acudir a la experimentación controlada, que no siempre se muestra pulcra en la selección de la muestra⁴². Sobran los ejemplos de comportamientos que —sin retorcer el sentido de las palabras— no cabe explicarlos con el supuesto de que las personas solo atienden a sus propios intereses: los constituyentes franceses del 4 de agosto de 1789 o los del 16 de mayo de 1791 que, “ebrios de desinterés”, renunciaban, respectivamente, a sus privilegios y a la posibilidad de ser elegidos en la asamblea legislativa⁴³, las muchas variantes de “misiones suicidas” (kamikazes, terroristas que “se autoinmolan”)⁴⁴ o la modesta

paradoja del voto, esto es, el hecho indiscutible de que los ciudadanos se molesten en participar (informarse, ir al colegio electoral, hacer cola, etc.) a pesar de que su voto personal, el único que controlan, nada puede decidir⁴⁵. Más en general, los muchos casos en los que la realidad se aleja de lo que según el dilema del prisionero debiera suceder, nos muestran que en las motivaciones humanas hay algo más que egoísmo: hay normas o emociones que funcionan con independencia de las posibles sanciones.

Lo que parece innegable una suerte de “instintos” morales o de justicia regulan nuestro comportamiento y, en no poca medida, parecen imponerse a cualquier cálculo. Una variante del llamado juego del dictador ejemplifica lo que quiero decir. En este juego, uno de los participantes —el “dictador”— impone el reparto de 100 euros con un compañero. En su versión original el dictador (propietario) decide lo que le da al otro, que puede ser desde todo a nada. En la nueva versión el dictador (sistema impositivo) no tiene nada al principio, pero puede fijar al otro jugador un “impuesto” sobre los 100 euros, en los mismos términos, desde todo a nada. Pues bien, cuando los participantes calificaron la generosidad de los dictadores, juzgaron que el grupo de los “dictadores impositivos” era más injusto que el de los “dictadores propietarios”, incluso aunque sus propuestas distributivas sean más igualitarias: un dictador que entregaba 50 euros a su compañero era tenido por más generoso que otro que solo exigía recibir 30 euros⁴⁶. Habida cuenta de que no hay razones para atribuir justicia (objetiva, desde la perspectiva de la tercera persona) a la distribución de propiedad inicial (los 100 euros no son el resultado del esfuerzo de nadie, sino un regalo, un producto social, si se quiere), parece claro que la valoración de la justicia (por parte de los individuos) de la situación depende de la descripción de “derechos” asociada a la presentación del juego, mejor dicho, a como es percibido por parte de los individuos. Sencillamente, la justicia nos importa. Por eso nos parece mal que A dispon-

ga del doble que B si esa situación es resultado de un robo y nos parece bien si es resultado de un mayor esfuerzo o de una incapacidad que reclama asignarle mayores recursos.

B. SEGOS, HEURÍSTICOS E INCOHERENCIAS

El HE procesa la información como si fuera un ordenador. Compara, pondera correctamente, asigna probabilidades a las distintas opciones y, atendiendo a sus preferencias y a lo que puede pasar, decide. Los seres humanos reales funcionamos de manera bien distinta. Nuestras capacidades cognitivas son el resultado de procesos evolutivos en los que importan muchas cosas a la vez, no solo “la verdad”. En ambientes complejos y cambiantes, con información limitada, con datos incompletos y parciales, con una memoria limitada, no estamos en condiciones de procesar con meticulosidad neurótica los datos sobre alimentos, territorios, presas o predadores. No toda esa información puede estar codificada en nuestro genoma. Tampoco sería conveniente: un cambio en el entorno nos dejaría a ciegas, con programas caducados operando a piñón fijo. Mejor estar abiertos ante los cambios. Abiertos no quiere decir vírgenes. La idea de la mente en blanco tampoco encaja en el relato evolutivo. Hay las suficientes constantes en el mundo, en las relaciones causales, en la continuidad de los objetos, en las necesidades de nutrición, como para qué resulte ventajoso haber sedimentado en nuestro cableado mental ciertas capacidades cognitivas⁴⁷.

En esas condiciones, cuando importa reaccionar con prontitud a partir de unos pocos indicios, sirve de mucho aplicar unas cuantas reglas que nos ayuden a simplificar las decisiones, aunque no sean muy pulcras metodológicamente⁴⁸. En la sabana, ante un movimiento extraño de la vegetación, “inferir” la presencia de un depredador y salir corriendo, no es un ejemplo de buen proceder científico. Lo debido sería repetir el experimento en circunstancias cambiantes antes de extraer conclusiones

precipitadas. Pero si lo que importa es salvar la vida, es mejor no andarse con demasiadas finuras y tirar por lo derecho.

Así las cosas, no nos podemos extrañar de que, en nuestros razonamientos reales, nuestra mente nos predisponga al uso regular de atajos mentales o heurísticos. Hemos aludido a alguno, pero hay muchos más. El de disponibilidad, que conduce a valorar las probabilidades a partir de la experiencia más a mano, explica que después de un accidente de avión disminuya el número de pasajeros o que los médicos que han diagnosticado dos casos seguidos de meningitis reconozcan los mismos síntomas en el próximo paciente, con independencia de la extensión real de la enfermedad en la población. El de confirmación, la predisposición a solo recoger argumentos y datos favorables a nuestras ideas e ignorar los que no se ajustan a ellas⁴⁹. El de ancla, según el cual la primera información condiciona las opiniones posteriores, como sucede con el primer precio de un producto que nos lleva a percibir como baratos (o caros) los posteriores. El de exposición que configura nuestras preferencias u opiniones según el orden en que aparezcan, como sucede cuando la secuencia de aparición de los alimentos en un comedor escolar decide la composición final del consumo o cuando no contestamos lo mismo si se nos pregunta si somos felices después de preguntarnos si nuestras relaciones sexuales son satisfactorias que si el orden de las preguntas se intercambia. El sesgo del jugador que lleva a pensar que sucesos aleatorios individuales están influidos por otros previos, como si el hecho de que hayan salido hasta ahora cara al tirar una moneda sea una razón para que salga cruz. El sesgo retrospectivo (o falacia del historiador) que lleva a pensar que lo que acabó siendo importante fue importante antes de llegar a serlo. El sesgo del falso consenso que conduce a pensar que las propias opiniones son ampliamente compartidas. El de arrastre que nos lleva a participar de una opinión porque los demás la comparten. El de *statu quo*, la disposición a sobrevalorar las situaciones existentes o estables en el tiempo,

como se deja ver en las resistencias vecinales a cambios urbanísticos claramente beneficiosos⁵⁰. Son sesgos que afectan a errores inferenciales, a la recogida de información y a la (errada) manera de procesarla y que nos impiden describirnos como sujetos bayesianos⁵¹.

En otros casos nos encontramos con problemas que afectan a la incoherencia en las preferencias. También aquí hay experiencias y experimentos. Estamos dispuestos a pagar diez euros al hijo de nuestro vecino, como mucho, por cortar la hierba de nuestro jardín, lo que parece indicar que apreciamos más once euros que la tarea, aunque si el vecino nos ofreciera once por realizar la misma tarea en su jardín, idéntico al nuestro, seguramente lo rechazaríamos: lo primero parece indicar que apreciamos más once euros que el trabajo; lo segundo, lo contrario. Si estamos en la cola de un cine y alguien se ofrece a comprar nuestro puesto por el doble de lo que cuesta la entrada seguramente no lo aceptaríamos aunque sí, repentinamente, duplicaran el precio de las entradas nos iríamos a casa: la elección es la misma pero, en un caso, preferimos la entrada a los diez euros y, en el otro, lo contrario. La cantidad de dinero que estamos dispuestos a pagar por una medicación que nos evite el desarrollo de una enfermedad a la que hemos estado expuestos con cierta probabilidad (1/10000) no es igual a la que demandamos por participar en una investigación de una vacuna en la que existe la misma probabilidad de que desarrollemos la enfermedad⁵².

Entre los experimentos, quizá el más interesante —por sus implicaciones para la justificación de los diseños institucionales— es el llamado efecto apropiación, según el cual, otorgamos valor a los objetos de nuestra propiedad por el simple hecho de que sean nuestros⁵³. En los experimentos, las personas reclaman un precio más alto por un bien que se les acaba de dar de lo que están dispuestos a pagar por él. El simple hecho de que “sea suyo” —aunque no lo hayan elegido o comprado— les lleva a valorarlo. No es suyo porque lo prefieren, sino que lo prefieren porque es suyo. No

son las preferencias las que deciden las elecciones (la compra y la propiedad) sino la propiedad la que conforma las preferencias. La implicación general no es desatendible: no podemos decir que la estructura de propiedad de una sociedad expresa la voluntad (las preferencias) de los ciudadanos puesto que es la estructura de propiedad la que decide la estructura de preferencias. El problema, obviamente, es que no hay intercambio de mercado sin atribución previa de derechos de propiedad.

C. SIMPLICIDAD

Ninguno de los comportamientos inventariados son los que cabría esperar del HE. No somos ni tan egoístas ni tan racionales. Otro modo de decir eso mismo es reconocer que somos bastante más complicados. Algo que nos recordó en 1977 Amartya Sen en un artículo muy citado con el elocuente título de “tontos racionales”⁵⁴. En lo esencial, sostenía que el HE, además de una irrealidad empírica, es un simple. Mejor dicho: que de tan simple es falso. A diferencia de lo que sucede en casi todas las otras especies animales⁵⁵, nuestra psicología no es plana: no se limita a la racionalidad de objetivos y creencias. Como nos reconocía la segunda idea de racionalidad, las personas tenemos pensamientos sobre los pensamientos de los demás y, también, pensamientos sobre nuestros pensamientos. Aunque no cabe descartar que otras especies pueden disponer de la capacidad para atribuir estados mentales⁵⁶, resulta más que improbable que tengan pensamientos sobre los pensamientos propios⁵⁷. Esta circunstancia, por ejemplo, nos hace emocionalmente sofisticados: para poder experimentar metaemociones hemos de ser conscientes —de designar con una palabra— la emoción básica. El miedo se puede experimentar sin saberlo, la vergüenza de sentir miedo requiere que sepamos que experimentamos miedo.

Esa capacidad nos abre un mundo. Queremos no querer ciertas cosas y queremos querer otras, tenemos metapreferencias, órdenes de preferencias entre órdenes de preferencias⁵⁸. El

HE ordena sus preferencias y procura lo que está en primer lugar. En su casa, por la tarde, si a Juan le apetece entregarse a la televisión antes que hacer deporte, se sentará frente a la televisión. Un ser humano real es más complicado. Muy bien pudiera suceder que, después de algunas lecturas, a Juan le dé por pensar no en qué hacer con sus tardes sino en qué hacer con su vida, y, una vez contestada la segunda pregunta, abordar la primera, qué debe hacer con sus tardes. Quizá se ha forjado cierta idea de lo que es una vida valiosa y quiere llevarla a cabo. Esa experiencia, tan humana, complica la descripción racionalista del HE. Para que encaje el guión del HE es fundamental que exista un único objetivo, el que se maximiza: dado cierto nivel de precios y de consumo, puedo ordenar los automóviles por su rapidez. Pero quizá, desde mi mejor yo, considero que debería apreciar otras cosas: mis metapreferencias me llevan a “querer querer” la seguridad y, entre unas cosas y otras, mi comportamiento pierde coherencia.

Es cierto que, al final, se decide y cabría la solución de integrar todo en un valor único⁵⁹. Pero en otras ocasiones, la complejidad y la coexistencia abre la puerta a luchas psicológicas entre motivaciones diferentes con una difícil traducción en el guión de HE. El psicoanálisis se puede entender como un modo tosco de reconocer esa circunstancia: el yo, el ello y el superyo entendidos como conjuntos coherentes cada uno de ellos de creencias y deseos que se expresan, sin embargo, como contradictorios agregadamente⁶⁰. En una experiencia más familiar, la simpatía, para expresarse con el lenguaje de Adam Smith, nos lleva a tener en cuenta no solo nuestros intereses, sino también los de nuestros seres queridos. Un padre puede hacer suyos los ordenes de preferencias de sus hijos y, al unirlos con los propios, y agregarlos, encontrarse con una suerte de paradoja de Condorcet intrapsíquica. Entiéndase, no es que se participe de una esquizofrenia, como que se tengan diversas jerarquías, cada una de las cuales puede ser internamente consistente, pero que reunidas, den pie a inconsistencias.

El verso de Walt Whitman, lo sintetizaba con claridad: “¿qué me estoy contradiciendo?, no importa, soy amplio, contengo multitudes”. Yo puedo preferir A a B y B a C, pero, en virtud de mis afectos o mis odios, también participar de otro orden de preferencias B a C y C a A. Ese otro orden de preferencias puede ser el de la persona que quiero y cuyos deseos hago míos, o, al revés, el orden de preferencias que detesta una persona que odio, lo que me lleva a mí a suscribirlo⁶¹.

Pero también hay otra vía en la que la consideración de los otros complica el modelo del HE. Y es que nuestro bienestar no es independiente del de los demás. Algo que no contempla el HE. A éste le trae sin cuidado como le vaya a los otros, para bien o para mal. Su función de utilidad es independiente de la de los demás. El bienestar de los demás no aumenta el suyo, como le sucede al altruista, pero tampoco lo disminuye, como le sucede al envidioso. Algo bastante raro, al menos si hacemos caso de las investigaciones empíricas de la economía de la felicidad que, entre otras cosas, muestran que nos produce más satisfacción un aumento de nuestro sueldo de un 10% que una aumento del 20% común, a nosotros y a nuestros compañeros de trabajo. Puede resultar deprimente pero es así: nuestro bienestar depende del de los demás, sobre todo, de los que son más parejos a nosotros, y no precisamente porque nos llene de dicha el suyo⁶³. Con cínica sinceridad lo explicaba Gore Vidal: “Para ver verdaderamente felices, no basta con tener éxito uno mismo...Deben fracasar los amigos”.

Y por ese cabo aparecen otros problemas clásicos de la racionalidad que tampoco tienen un fácil acomodo con la idea de HE, en tanto complican la equiparación urgente entre preferencias, intereses y bienestar: las preferencias adaptativas y las contradaptativas. Las primeras se producen cuando los individuos optan por reajustar sus deseos a sus posibilidades, pero no al modo del budista, que no quiere lo que no le resulta accesible, pero no por ello deja de valorarlo, sino como la zorra que se miente a sí

misma diciendo que están verdes las uvas maduras que escapan a su alcance⁶³ o de las mujeres de la India que dicen estar contentas con su carencia de elementales derechos⁶⁴. Las segundas operan en la dirección contraria. Son las de aquellos para quienes siempre está más verde el jardín del vecino o esos otros convencidos de que “no hay otros paraísos que los paraísos perdidos”. Estas situaciones, tan comunes como para participar de tales avales científicos y literarios, complican buena parte de los soportes sobre los que se levanta el HE. Ya no podemos decir con tranquilidad que los agentes en sus acciones maximizan la utilidad.

Ni somos tan simples como el HE ni la complejidad de nuestro comportamiento es susceptible de reducción al cálculo racional e interesado. Los humanos presentamos procesos emocionales diversos, repertorios de compromisos flexibles, variedad de vínculos entre normas y emociones, relaciones diferentes entre emociones y toma de decisiones⁶⁵. En determinados escenarios actuamos según normas de intercambio, en otros, según principios comunitarios y en otras ocasiones predominan relaciones jerárquicas⁶⁶. Pero no solo es que en un mismo individuo coexistan distintos repertorios emocionales y normativos, que se activen en distintos escenarios, sino que también en un mismo escenario pueden coexistir individuos motivacionalmente bien diferentes⁶⁷. Vale la pena examinar esto aparte.

D. INCAPACIDAD SOCIAL

Los resultados descritos muestran la importancia en nuestras actuaciones de normas o sentimientos de justicia. De su hondura no hay quizá mejor prueba que la propia existencia del mercado, imposible sin una red moral y emocional. El mercado opera sobre un escenario normativo e institucional que hace que los acuerdos se cumplan, los intercambios legítimos se reconozcan, la propiedad se acepte, ciertos criterios de distribución se consideren justos y algunas formas de competencia se juzguen inconvenientes. Pero ese escenario, el

compromiso con las normas de lealtad o confianza, no se sostiene ni se puede sostener en el interés. Ningún HE empleará tiempo en cultivar normas que beneficiaran a todos sin exclusión. No tiene interés en asumir el coste de mantener la trama moral, pero sí tiene interés en que exista. Lo mejor, de hecho, es que los demás la respeten, pero no uno mismo. Desde cierta perspectiva, la trama moral funciona como un bien público. A todos nos beneficia su existencia, pero nadie está interesado en producirla porque no hay modo de venderla entre sus consumidores: si en la sociedad se “respira” confianza, no hay modo de excluir a nadie de su uso y cobra el consumo del bien.

Eso sí, si la conducta calculadora se extiende, los valores y sus beneficios desaparecen. El egoísmo, la mejor estrategia para cada cual, es la peor para todos. Dicho de otro modo, el mercado no produce la geografía moral que necesita para funcionar. En realidad, alienta comportamientos que minan esa geografía: como vimos, el HE hace lo que no puede dejar de hacer⁶⁸. Y sin embargo, la trama persiste. En cierto modo, que el mercado sobreviva cuando atenta contra las condiciones mismas de su funcionamiento, que persistan conductas que sistemáticamente se ven penalizadas, es en sí mismo, una invitación a pensar que las disposiciones morales tienen raíces más profundas que la pura convención cultural o el simple acuerdo de intereses⁶⁹.

Raíces que no proceden de la racionalidad. El vínculo calculado con las normas las desactiva. La red moral solo puede cumplir su función si no es resultado de una elección estratégica, si no se sustenta, ella misma, en el interés. Las normas resultan “interesantes”, para el orden social, pero no pueden escogerse o seguirse por interés. Resultan beneficiosas si el compromiso con ellas no obedecen a un cálculo, si no se actúa por las ventajas que reportan, solo si “se sienten” genuinamente. Si en una negociación en mitad de una industria común, tú, que tienes por el mango la sartén del reparto del producto, me ofreces elegir ente una miseria o nada, en buena lógica económica, yo

debería escoger la miseria. Es lo que haría el egoísta racional. Algo es mejor que nada. Un euro es mejor que cero. Pero si sabes que no tolero el trato indigno, te lo pensarás dos veces antes de humillarme. El poderoso no prolonga las negociaciones *ad eternum* porque anticipa que sentimientos de justicia impiden a las personas aceptar acuerdos indignos, aunque la alternativa sea perderlo todo. Es la enseñanza del juego del ultimatum. Si el orgullo o la vanidad herida me ciegan y prefiero morir de hambre a sentirme humillado y tú conoces ese cerrilismo mío, muy probablemente mi irracionalidad te llevará a ofrecerme algo más que la miseria, ante el temor de que la empresa común se acabe cuando yo haya acabado conmigo y con todo: si me marchó y rompo la baraja, no habrá cooperación y todos perderemos.

Para eso, para que tu sientas mi obstinación como una amenaza, es importante que mi sinrazón no sea fingida, que no sea un farol estratégico, que dura un instante y que, una vez el dinero en la mesa, ante el *fait accompli*, cuando ya no quedé sino elegir entre algo o nada, acabaré por plegarme. Si yo, en el fondo, ando con cálculos y tú sospechas que, a la hora de la verdad, actuaré racionalmente, que preferiré algo a nada, no cederás a lo que no deja de ser una presión retórica que sabes simulada. A la larga la conducta ciega, honesta o rencorosa, emocional, “sale a cuenta” tanto para el individuo como para la propia sociedad, para sus mecanismos reproductivos. El buen funcionamiento de la economía y hasta la consecución de los mejores intereses de las personas con frecuencia exigen que éstas se olviden de sus intereses más inmediatos y hasta de la racionalidad. La vida social, incluida la económica, resulta imposible entre *homines economici*.

En ese contexto, no solo es importante que los sujetos se dejen llevar por las emociones, que no actúen porque les “sale a cuenta”, aunque les salga a cuenta, sino, sobre todo, que sean capaces de interpretar mutuamente sus emociones. La dignidad solo sirve si es sentida e irrenunciable y si el otro, que es como yo, entiende las señales. Si los valores “funcionan”

y aseguran la vida social, es precisamente porque no se asientan en el cálculo, porque son normas, sostenidas en emociones y, al fin, en nuestra historia evolutiva. La selección natural es el mecanismo que nos permite dar cuenta de la persistencia de comportamientos y disposiciones (morales) que “salen a cuenta” para todos pero que no cabe entender desde el beneficio de cada uno⁷⁰. Las emociones, el instinto y la biología. Sin esa trama no hay indignación ante la injusticia, ante la violación de normas, y sin normas no hay economía. En breve, una sociedad de egoístas racionales resulta socialmente imposible.

En realidad, no parecía descaminado Sen. El HE, en su simplicidad, está incapacitado para la vida compartida. Desde cierto punto de vista el HE se parece, en más de un sentido, a esos idiotas morales y emocionales que han estudiado los neurobiólogos: individuos con ceguera afectiva, incapaces de tomar decisiones o de valorar el significado emocional de los acontecimientos⁷¹. Tales sujetos, o los monos sociales a los que se les cortan conexiones cerebrales que reproducen sus lesiones, resultan personajes desnortados socialmente. Carecen de emociones y no “sienten” las normas, incluso si las siguen por convención, como los demás seguimos las normas de etiqueta: la injusticia o la vejación no les indignan, si acaso, les parecen una inconveniencia, una grosería. Se muestran incapaces de valerse en sociedad: para competir o para cooperar o para conocer su ubicación social, para tratar con los demás, para percibir el significado emocional (las inflexiones) de las distintas palabras. La falta de emociones morales (de justicia o de empatía) les incapacita para calibrar el sentido de los conflictos, para sentir la culpa y la vergüenza o el agravio que acompañan a procesos sociales y que cumplen la función de coordinar las conductas. Todos sabemos cómo interpretar y valorar que tú me arrebatas la comida: te estigmatizaremos y, tú, avergonzado, procurarás no hacer ruido. La emoción de ira o enfado de un miembro del grupo se complementa con la vergüenza o culpa que experimenta el que ha

realizado una mala acción. Este último acepta la sanción sin que el conflicto cuaje en una violencia que haría desaparecer el grupo. La comunidad sobrevive merced al reconocimiento de que se ha violado una norma, la emoción correspondiente y el conocimiento por todos de ambas circunstancias. Todo ello al HE le suena a chino⁷².

5. CONCLUSIÓN: UNA VALORACIÓN MATIZADA

El repaso anterior invitaría a olvidarse del HE. La conducta de los humanos reales se aleja de la conducta del HE y, además, los desvíos son regulares, lo suficiente como para que podamos identificar la existencia de unos sesgos. El reconocimiento de esta circunstancia ha llevado a más de uno a decretar la muerte del personaje⁷³. Los más moderados, recomiendan cultivar una teoría más realista sostenida en “un modelo de actores humanos enriquecido (empíricamente)”⁷⁴. En esa dirección se incluirían aquellos que reclaman debilitar las exigencias de racionalidad: más que buscar el óptimo los seres humanos nos contentamos con un suficiente, sea por nuestras limitadas capacidades cognitivas sea porque la idea de obtener el mejor resultado, en escenarios cambiantes, se revela un imposible, algo así con disparar sobre un blanco móvil, como si cambiara la competición cada vez que aprendemos las reglas del juego⁷⁵. En otros casos se ha optado por abandonar la hipótesis del egoísmo en aras de otras motivaciones: unos sostienen que el humano real apuesta antes por la reciprocidad que por el egoísmo⁷⁶ y otros que, en realidad, no nos regimos por un único motivo o principio, sino por distintos, según el contexto o las circunstancias: el mercado, la familia, el trabajo⁷⁷.

Las descalificaciones más radicales proceden de quienes sugieren abandonar el guión completo de R, el supuesto de agentes intencionales. A esa conclusión llegan algunos que sostienen que creencias y deseos asumen una

ontología incompatible con lo que conocemos, que “ha permanecido estancada por veinticinco siglos, y cuyo marco categorial se ha mostrado inconmensurable con las categorías de la ciencia física”⁷⁸. R puede resultar útil para ir por el mundo y entendernos, pero no cabría la posibilidad de relacionarlo conceptualmente con nuestras teorías sobre el cerebro humano, al modo como la biología molecular se relaciona —se reduce a— con la química y ésta con la física. En ese sentido se trataría de una teoría psicológica, pero una teoría falsa, como la del flogisto o la del calórico o como la física aristotélica, esa que nos lleva a decir que “el Sol sale”: nos entendemos, podemos anticipar acontecimientos (“mañana saldrá al Sol”), útil pero irreal⁷⁹. Con la misma intención descalificadora del guión completo, otros dirán que, en el fondo, los verdaderos protagonistas de nuestros comportamientos son los genes y que el resto, todas nuestras metas humanas (el amor, la religión, los valores) no son más que los subproductos de un cerebro predispuesto a elaborar justificaciones⁸⁰. Nosotros seríamos poco más que vehículos reproductores, portadores de genes, viviendo en un mundo de ficción. Nuestras acciones no responderían a decisión alguna, por más que, una vez realizadas, generen una necesidad de justificación, tarea de la que, según los neurobiólogos, se encargaría nuestro hemisferio cerebral izquierdo dedicado a interpretar —a proporcionar relatos de— nuestros comportamientos⁸¹.

La solvencia de estas conjeturas es dispar. No cabe desatender su vigor crítico, pero, por el momento, en su mayoría, no pasan de ser especulaciones⁸², compatibles con muchos resultados, pero, en sí mismas, no sobradas de avales empíricos y, desde luego, incapaces de proporcionar desarrollos como la teoría social sostenida en torno al HE⁸³. Por supuesto, que dispongamos de desarrollos no quiere decir que debamos apostar por ellos, si resultan inútiles. Debemos evitar la respuesta que el premio Nobel de economía, Kenneth Arrow, observó entre los militares, cuando oficiaba como meteorólogo durante la segunda guerra mun-

dial: “Los estadísticos que había entre nosotros verificaron las previsiones y descubrieron que no diferían del azar. Los propios encargados de preparar los pronósticos estaban convencidos de tal extremo y pidieron que dejaran de hacerse. La respuesta decía aproximadamente lo siguiente: “El general en jefe es consciente de la inutilidad de las previsiones meteorológicas. Sin embargo, las necesita por motivos de planificación”⁸⁴.

Ciertamente, las investigaciones examinadas han confirmado el desvarío del imperia- lismo de la economía forjado en torno a un modo de persona egoísta e hiperracional. Ese resultado ya nadie lo discutirá. Ahora bien, eso no quiere decir que el HE resulte una pieza inútil desde el punto de vista teórico ni que podamos descartarlo para siempre. De modo que, aunque solo sea para sopesar exactamente su alcance, y el de las críticas, quizá conviene terminar con una notas acerca de cómo no deben interpretarse los problemas expuestas.

1. Que no seamos egoístas en muchas ocasiones no quiere decir que seamos altruistas a tiempo completo. Los abundantes experimentos destinados a refutar que el interés material sea nuestra motivación exclusiva hay que interpretarlos en su exacto sentido: no es verdad que en todo momento y circunstancia seamos *homine economicii*. Pero eso no quiere decir que el HE resulte incondicionalmente falso. La dinámica de Newton, verdadera en ciertos sistemas, deja de servir cuando nos aproximamos a la velocidad de la luz, la mecánica clásica deja de valer para dimensiones cercanas a las constante de Planck y la teoría de la selección natural, capaz de explicar la evolución de las especies, no da cuenta de la evolución del sistema solar o de las sociedades humanas. En el mismo sentido, podemos reconocer que el HE, falso en un convento, una comuna o una familia, funciona en muchos sistemas reales, como lo confirman cada día millones de decisiones ante variaciones en los tipos de interés, la bolsa, los mercados cambiarios, la compra venta de pisos o las lonjas de pescado. Al igual que el “experimento” de que un trabajador, ante dos

ofertas de empleo idénticas, opte por la mejor retribuida —y los otros miles que se podrían diseñar en la misma dirección— no refuta, sin más, que seamos altruistas, la experiencia de que, en muchos casos, no actuamos por dinero no refuta incondicionalmente el egoísmo, aunque sí la idea del egoísmo universal.

2. Que no seamos racionales no quiere decir que no podamos ser racionales. En indiscutible que los sesgos descritos son frecuentes y sistemáticos. Ahora bien, eso no nos impide reconocerlos y corregirlos. Del mismo modo que la noción de ilusión visual solo resulta inteligible sobre el trasfondo de la imagen correcta, la noción de sesgo tiene sentido bajo el contraste de la buena inferencia. Uno no puede perderse si no hay un camino correcto. Por lo demás, es una experiencia cotidiana el reconocimiento de que podemos escapar y corregir nuestras limitaciones, constricciones y disposiciones, incluidas las asentadas en nuestra biología. Sabemos que hay longitudes de onda que no percibimos, que tenemos una disposición natural a que nos gusten los alimentos dulces y que participamos de “instintos” agresivos o posesivos. Hemos elaborado teorías sobre esas cosas y conocemos la existencia de tales longitudes de onda, somos capaces de evitar los dulces y valoramos nuestros “instintos” y nos sobreponemos a ellos en nombre de otras consideraciones. Con la ayuda de la ciencia o de la reflexión moral, podemos reconocer esas determinaciones de origen, incluidas las más fundamentales y “deterministas”, y actuar en consecuencia⁸⁵.

3. Un corolario de lo que se acaba de decir es que, aunque la teoría de la racionalidad, en su versión más depurada, como teoría de la decisión, resulte falsa descriptivamente, eso no la invalida normativamente. En nuestras decisiones cotidianas, los seres humanos no comparamos y evaluamos con plena pulcritud analítica. Tampoco respetamos impecablemente, al formar nuestras creencias, los axiomas de la teoría de la probabilidad. Ni nosotros, ni puestos a decirlo todo, los teóricos de la estadística. En ese sentido, como teoría descrip-

tiva, en su versión más refinada, el HE sirve de poco. Pero esto no descarta la teoría de la racionalidad como ideal regulador, como base de las decisiones meditadas. Que cometamos errores al formar nuestras creencias no invalida el reconocimiento, por ejemplo, de que el cálculo de probabilidades —el conjunto de reglas no-inductivas de la lógica bayesiana— es la representación perfecta de la coherencia al formar las creencias. Pudiera suceder con la teoría de la racionalidad algo parecido a lo que sucede con buena parte de la teoría económica sobre los mercados perfectos, que, aunque esta pensada en origen para explicar el mercado, también actúa como un ideal regulativo de las acciones: nos señala un norte, construir en la realidad las condiciones que permiten los buenos resultados⁸⁶.

4. Que las emociones y los instintos morales nos ayuden a decidir no los convierte en principios de racionalidad práctica. No son la última palabra. En ningún caso suplen a la argumentación, científica o moral. Las decisiones instintivas muchas veces aciertan, pero, para saberlo, hemos de poder aquilatarlas con los mejores procedimientos, con la razón y la experiencia⁸⁷. Las emociones no resuelven los dilemas morales, no nos dicen que está bien o mal ni, por supuesto, están más allá de nuestra valoración, de la posibilidad de ser juzgadas. Aunque nos ayuden a decidir y valorar, no son las que, al final, valoran sino las valoradas. Nuestras emociones existen porque cumplen y, sobre todo, porque cumplieron funciones adaptativas en los entornos en los que se ha desarrollado la mayor parte de la vida de la especie, entornos por lo general muy alejados de aquellos por los que transitamos en el presente. Algunas de esas emociones que hoy nos disgustan se asentaron en nuestro cableado mental por su provecho en otro tiempo, eran interesantes dada nuestra condición de cazadores recolectores (por eso somos agresivos), para asegurar nuestra particular herencia genética (por eso somos celosos) o para prevenirnos frente a otros grupos cuando hay pocos recursos (por eso somos racistas). Ahora bien, su persisten-

cia, indiscutible, no impide que condenemos y castigemos los comportamientos violentos o sexistas desencadenados por ellas. Nos sucede lo mismo que con nuestra disposición hacia los alimentos dulces, con alto contenido energético, de mucho interés en los entornos naturales en los que ha transcurrido la vida de nuestra especie, con pocos recursos: hoy, en la sociedad de la abundancia, valoramos esa disposición, atendiendo a nuestra salud, la juzgamos inconveniente y evitamos atender a sus demandas. Vale la pena dedicar un nota más al trasfondo donde arraiga este potencial normativo, el soporte de la racionalidad.

5. Nuestra capacidad de abstracción y, por ende, la posibilidad de tasar emociones e intuiciones, es inseparable de un talento específicamente humano, también anclado en la biología (otra cosa, discutida, es si como resultado adaptativo de la selección natural)⁸⁸ que amplía nuestro mundo y nos permite escapar a nuestras constricciones biológicas: el lenguaje. La complejidad del lenguaje humano lo otorga un enorme potencial liberador. Su radical novedad respecto a cualquier otra forma de comunicación animal se muestra en sus dos componentes centrales: el uso extensivo de arbitrarias conexiones (aprendidas) significado-forma y su combinación en sistemas sintácticos⁸⁹. A partir de ahí se desplegaron ciertas posibilidades: recursividad, abstracción, combinatoria, potencial simbólico, etc. Ese paso es el cimiento en el que se basa nuestra posibilidad de revisar emociones, sesgos y disposiciones. Las nuestras y las de los demás, la discusión y la cooperación⁹⁰. Nos abre el universo de lo que no es pero puede llegar a ser, el universo desde donde mirarnos y corregirnos. Elaboramos pensamientos complejos, corregimos juicios y comparamos argumentos que recaen sobre sociedades hipotéticas, experiencias pasadas, instituciones inexistentes o soluciones a problemas que todavía no se han dado. Todas actividades son posibles por el lenguaje y, entre ellas, destaca nuestra capacidad para construir contrafácticos, conjeturas acerca de qué pasaría si en lugar de A se hubiera dado B. Esa capa-

cidad de pensar la realidad sobre el horizonte de sus alternativas está en el origen de nuestras explicaciones de lo que sucede, de cómo es el mundo, y también de nuestros anhelos, de nuestras esperanzas acerca de cómo nos gustaría que fuese. Esa simple posibilidad, de revisar y rectificar nuestros juicios a la luz de razones, confirma nuestra competencia racional y nuestra capacidad para desarrollarla⁹¹.

Este último punto, que ciertamente nos aleja del HE, tampoco nos acerca a ese nuevo naturalismo que algunos parecen inferir de los resultados descritos y que se traduce en una suerte de entrega, como horizonte de posibilidades, a nuestras rutinas de razonamiento, emociones o instintos, a la irracionalidad, para decirlo en una palabra. También estos pueden ser valorados. Ese es el más genuino talento de la especie humana, su capacidad natural para emanciparse de su naturaleza.

NOTAS

1. Un repaso de su historia y sus limitaciones, cf., P. Demeulenaere, *Homo oeconomicus: Enquete sur la constitution d'un paradigme*, París, PUF, 1996.
2. A. Hirschman, *The Passions and the Interests: Political Arguments For Capitalism Before Its Triumph*. Princeton, Princeton U. P., 1977.
3. En opinión de Montesquieu, *De l'esprit des lois*, [1748] en *Oeuvres Complètes*, París, Seuil, 1964, pág. 546.
4. Para un repaso en el marco de la teoría económica, hasta recalcar en las teorías del equilibrio general, cf. B. Ingrao, G. Israel, *The Invisible Hand. Economic Equilibrium in the History of Science*, Cambridge, Mass., The MIT Press, 1990.
5. G. Vico, *Principi di Scienza Nuova*, L'Aquila, REA Edizioni, 2013 (e.o. 1725) libro I, II, VII. Kindle e-book
6. J. Steuart, *Inquiry into the Principles of Political Economy*, Chicago, University of Chicago Press, 1966 (e.o. 1767), págs. 143-144.
7. La idea será particularmente desarrollada por la escuela austriaca de economistas (Von Mises, Hayek). Se la denominará "Catalaxia", cf. M. Rothbard, "Catalactics" en J. Eatwell, M. Milgate, P. Newman (ed.), *The New Palgrave: A Dictionary of Economics*, Vol. 1, Hampshire, Palgrave Macmillan, 1987, págs. 377-378.
8. J. Diamond, *Colapso: por qué unas sociedades perduran y otras desaparecen*, Madrid, Debate, 2005.
9. Algo parecido sucede con la extensión de las lenguas o los sistemas de medidas. Se dan aquí diversos procesos superpuestos: economías (o externalidades) de red, dependencia de la senda: M. Katz, C. Shapiro, "Systems competition and network effects", *Journal of Economic Perspectives*, 1994, 8; P. Garrouste, S. Ioannides (eds), *Evolution and Path Dependence in Economic Ideas: Past and Present*, Cheltenham, Edward Elgar Publishing, 2000. Todos ellos se pueden interpretar como variantes del mecanismo de la mano invisible. Para el caso de la lengua, R. Keller, *In Language Chance*, Londres, Routledge, 1994.
10. A. Sen, "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", en A. Sen, *Choice, Welfare and Measurement*, Oxford, Blackwell, 1982, pág. 89.
11. Hay un cierto paralelismo con el proceder de las versiones más primitivas de la selección natural, con su interdependencia entre las ideas de adaptabilidad y la de supervivencia: la tesis de que solo sobreviven los más aptos se vaciaba de contenido en el mismo momento en que la aptitud se definía como la capacidad para sobrevivir.
12. A. Rosenberg, "Philosophy of Social Science", W. Newton-Smith (ed.), *A Companion to the Philosophy of Science*, Oxford, Blackwell, 2000. cf. también A. Rosenberg, "What is the Cognitive Status of Economic Theory", en R. Backhouse (ed.), *New Directions in Economic Methodology*, Londres, Routledge, 1994.

13. J. Speth, “las estrategias alimentarias de los cazadores recolectores”, *Mundo científico*, 1987, 73.
14. D. Dennet, *The Intentional Stance*, Harvard, The MIT Press, 1989, pág. 265.
15. D. Dennet, *Kinds of Minds*, Londres, Weindenfeld, Nicholson, 1996, pág. 27.
16. En un sentido parecido al que podemos decir que $y = mx + n$, mientras no se interpreten las variables (x, y) como variables numéricas de propiedades de un sistema real, no es propiamente una (supuesta) ley científica, sino un esquema de ley: todo esquema (de ley, en este caso) puede recibir una infinidad potencial de interpretaciones fácticas, cf. M. Bunge, *La investigación científica*, Barcelona, Ariel, 1983, pág. 340.
17. Algunos (Sócrates R. M. Hare) dirían que los comportamientos de los débiles de voluntad, *akráticos*, muestran que sí es posible: el individuo que, a pesar de tener ciertos objetivos (dejar de fumar, terminar una relación, etc.) y buena información, no actúa de acuerdo con sus creencias y sus deseos. Otros dirán que, al final, también el akrático es racional: si hace X es porque quiere X. La discusión, su posibilidad de solución, si es que la tiene, depende de la idea que se tiene de R, si tiene una naturaleza teórica, metodológica o —incluso— lógica, cf. S. Stroud, C. Tappolet, (eds.), *Weakness of Will and Practical Irrationality*, Oxford, Clarendon Press, 2003.
18. Aunque se podría optar por una enmienda general: si por ejemplo fuéramos una suerte de zombies preprogramados (en cuyo programa se incluiría la ilusión de actuar por creencias y razones). Volveré al final sobre diversas conjeturas que apuestan por prescindir del guión de la racionalidad, cf. F. Ovejero, “Economía y psicología: entre el método y la teoría”, *Revista Internacional de Sociología*, 2004, 38.
19. L. Cosmides, J. Tooby, “Beyond Intuition and instinctive blindness: Toward an evolutionary rigorous cognitive science”, *Cognition*, 1994, 50, pág. 64.
20. H. Plotkin, *Evolution in Mind*, Londres, Alen Lane, 1997, pág. 139.
21. En las que se participa del mismo “entorno cognitivo”: los individuos “comparten información sobre la información que comparten”, D. Sperber, D. Wilson, *Relevance* (2ª ed.), Oxford, Blackwell, 1995, pág. 45. Fórmula que es un desarrollo de la idea de “conocimiento común (mutuo)” de D. Lewis, *Convention*, Cambridge, Harvard U.P., 1969. Resulta interesante destacar que una idea parecida (“the tacit cognitive communalities”), “entre el instinto y la razón” es formulada desde la teoría de la elección racional para referirse a los contextos normativos, U. Witt, “Moral norms and rationality within populations: an evolutionary theory”, J. Casas Pardo, F. Schneider (eds.), *Current Issues in Public Choice*, Cheltenham, Edward Elgar, 1995, pág. 251.
22. D. Sperber, D. Wilson, op. cit. págs. 65-22.
23. Para una breve y sugestiva exposición del vínculo evolutivo entre los humanos como “psicólogos naturales” y la vida social, cf. N. Humphrey, *La mirada interior*, Madrid, Alianza, 1993. Precisamente por esas posibilidades estratégicas de inducir en los otros creencias falsas —resultado al fin del reto de hacer compatibles la competencia y la cooperación— a este talento se lo califica de inteligencia maquiavélica, cf. Richard W. Byrne y Andrew Whiten (eds.), *Machiavellian Intelligence*, I. *Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, Oxford, Oxford U.P., 1988; Richard W. Byrne y Andrew Whiten (eds.), *Machiavellian Intelligence*, II. *Evaluations and Extensions* Oxford, Oxford U.P., 1997.
24. Repárese en que esta capacidad es condición de posibilidad de un supuesto clásico de la teoría moral: reconocer a otro como racional supone un principio de respeto y preocupación (y a partir de ahí, otro supuesto: la tesis de dar razones que los otros puedan aceptar), cf. T. Nagel, *The Possibility of Altruism*, Princeton, Princeton U.P., 1970; T. Scanlon, *What We Owe to Each Other*, Cambridge, Mass., Harvard U.P., 1998. Con más provisionalidad, aquí también se podría anclar una crítica a la tesis utilitarista que toma el sufrimiento como criterio básico de pertenencia a la comunidad de valoración, criterio que permite incluir sin matices a los animales no humanos. Para una teoría normativa atenta a los resultados cognitivos, la capacidad de sufrimiento no es independiente de la existencia de unos sistemas articulados de deseos, expectativas, de unos estados mentales complejos y con capacidad discriminatoria, que no se dan por igual en todas las especies. Cf. D. Dennet, “Animal Consciousness: What Matters and Why”, *Social Research: In the Company of Animals*, 1995, 62, 3, págs. 691-710
25. Hay un fuerte vínculo entre racionalidad, interacción social e identificación de los “mentirosos”, L. Cosmides, “The Logic of social exchange: has natural selection shaped how humans reasons?”, *Cognition*, 1989, 31.

26. J. Tooby, L. Cosmides, "Friendship and the Banker's Paradox: Other Pathways to the Evolution of Adaptations for Altruism", W. Runciman, J. Maynard Smith, R. Dunbar (eds.), *Evolution of Social Behavior Patterns in Primates and Man*, Oxford, Oxford U.P., 1996, págs. 129-130.

27. De no acabar de deslindar entre estos planos arranca, a mi parecer, la tesis de la escuela austriaca, según la cual, la economía es una ciencia *a priori*, como la lógica, la praxeología de Ludwig von Mises, *La Acción Humana: Tratado de Economía*, Madrid, Unión Editorial, 2011 (e.o. 1949). Los austriacos, eso sí, juegan fuerte. Para ellos son "sintéticas y a priori" bastantes más cosas que la racionalidad: la ley de la utilidad marginal decreciente, la ricardiana de asociación, la teoría cuantitativa del dinero, etc.

28. Se podría decir que esta idea es una teoría sobre (el contenido de) la idea anterior: asume que somos agentes racionales, que tenemos ese talento, y lo precisa en un sentido específico.

29. G. Becker, *The Economic Approach to Human Behavior*, Chicago, The University of Chicago Press, 1976, pág. 14.

30. R. Frank, *Microeconomics and Behavior*, McGraw-Hill, N. York, 1991, pág. 17.

31. A. Sen, *Rationality and Freedom*, Londres, Harvard U. P, 1990, pág. 202. Idea que se asocia a la "posibilidad de explicar el conjunto de elecciones reales como resultado de la maximización según una relación binaria" y que en las formulaciones más exigentes se entiende como "completamente transitiva o, en las más exigentes todavía, como representable mediante una función numérica que la persona pueda describirse como maximizando", A. Sen, *On Ethics and Economics*, Blackwell, Oxford, 1987, págs. 12-13.

32. A. Sen, *Rationality and Freedom*, op. cit., pág. 19, cf. asimismo, A. Sen, "Rational Behavior", J. Eatwell, M. Milgate, P. Newman (eds.), *Utility and Probability. The New Palgrave*, Londres, Macmillan, 1990.

33. En este contexto la distinción entre creencias y escenarios resulta bastante convencional. En los escenarios con certidumbre estamos en condiciones de saber qué sucederá (en este banco me dan el 3%, en el otro 2%, etc.), en los de riesgo podemos asignar probabilidades (20% de que llueva, etc.), en los de incertidumbre no tenemos ni idea. La distinción entre los dos últimos casos se podría discutir. En el caso de la incertidumbre la acción racional lleva a utilizar el criterio de maximización de la utilidad esperada basada en probabilidades subjetivas (bayesianas, en su versión personalista), esto es, los agentes, aunque no conozcan por anticipado las probabilidades de los distintos escenarios, siempre están en condiciones de asignar probabilidades que miden su grado de creencia, aplicando ciertos teoremas de la teoría de la probabilidad. Visto así, se puede decir que los escenarios de incertidumbre desaparecen porque los agentes siempre tienen una base probabilística para decidir. En todo caso, lo relevante para nosotros es que la teoría sostiene que las creencias del HE satisfacen los axiomas de la teoría de la probabilidad.

34. Pues muy bien puede suceder que se tengan las creencias y las preferencias, que se produzca la acción acorde con las creencias y las preferencias, pero que la acción no se explique por ellas. Con un ejemplo de Davidson: después de un accidente, un alpinista (B), odiado por su compañero de cordada (A), se queda suspendido sin otro amarre de la fuerza de A, quién sabe que si abre su mano el otro se matará, cosa que desea; finalmente, se le acaba por soltar la mano por la fatiga: tenía las creencias, las preferencias, se produjo la acción, pero la acción no se desencadenó por las creencias ni los deseos. De todos modos, como me observa el anónimo revisor de este trabajo, se podría discutir el ejemplo: "soltar la mano por fatiga" no es, en rigor, una acción y ni "la acción" es la misma respecto a las preferencias iniciales.

35. Hay aquí un asunto interesante no exento de problemas. Se puede actuar de acuerdo con una X sin que X sea la razón de la actuación. Nuestro comportamiento cotidiano (andar, interceptar un objeto, evitar otro, mantener el equilibrio, etc.) se pueden —en principio— describir a partir de una serie de funciones, sin que ello quiera decir que seamos conocedores de tales funciones, que estarían, por así decir, "insertas" en nuestro programa. En un sentido parecido, los economistas dirán el modelo lo que afirma es que "los campesinos iletrados, que no tienen ni idea de lo que es una función, y aún menos de lo que es una función cóncava, tienen preferencias representables por una función de utilidad cóncava, y resuelven cotidianamente los problemas de optimización (de un consumidor)", I. Ekeland, "Mathematiques et économie, fécondité et limites d'une formatisation", en I. Ekeland y J. Elster, *Théorie économique et rationalité*, París, Vuibert, 2011, págs. 9-10.

36. El supuesto, en principio, no es imprescindible, pero sin él, como dice Sen, "la teoría no va muy allá". De ahí que "no solo sea ampliamente usado en economía, sino que muchos de los teoremas centrales de la economía moderna (...) dependen de él", A. Sen, *Rationality and Freedom*, op. cit. págs. 20 y 23.

37. Para una exposición exhaustiva de cada uno de los componentes y del alcance de las críticas empíricas, aunque no estrictamente coincidente —en realidad crítica— con la de Sen, D. Hausman, *Preference, Value, Choice, and Welfare*, Cambridge, Cambridge U.P., 2011.
38. Por supuesto, siempre cabría parchear la racionalidad de HE y sostener, por ejemplo, que lo que comparamos al fin son utilidades y la utilidad de una moneda no es la misma que la de un billete del mismo valor, como lo “demuestra” que nos desprendemos con más facilidad de la moneda. A partir de ahí solo nos queda elaborar el conveniente relato ad-hoc (el peso de la moneda es un lastre, los billetes pueden servir en caso de necesidad para liarse un cigarro o en un lavabo, etc.).
39. Un repaso minucioso, crítico y con punto de vista propio de las dos dimensiones en J. Elster, *Le désintéressement: Traité critique de l'homme économique*, vol. I, París, Seuil, 2009; J. Elster, *L'irrationalité: Traité critique de l'homme économique*, vol. II, París, Seuil 2010.
40. W. Guth, R. Schmittberger, B. Schwarze, “An Experimental Analysis of Ultimatum Bargaining”, *Journal of Economic Behavior and Organization*, 1982, 3.
41. Para experimentos en esta línea, J. Kagel, A. Roth, *The Handbook of Experimental Economics*, Princeton, Princeton U.P. 1995. Sobre el “homo reciprocans”, S. Bowles, H. Gintis, “Homo Reciprocans”, *Nature*, 405, 2002; H. Gintis, J. Henrich, S. Bowles, R. Boyd, E. Fehr, “Strong Reciprocity and the Roots of Human Morality”, *Social Justice Research*, 2008, 21, 2, págs. 241-253.
42. Es común el reproche de la baja calidad representativa las poblaciones (por ejemplo, estudiantes de universidades de élite que acuden voluntariamente a los experimentos). Más en general, sobre los problemas metodológicos, D. Mayo, “Some Methodological Issues in Experimental Economics”, *Philosophy of Science*, 2008, 75. Un panorama exhaustivo: N. Bardsley, R. Cubitt, G. Loomes, et al., (eds.), *Experimental Economics: Rethinking the Rules*, Princeton, Princeton U.P., 2010.
43. J. Elster, *Le désintéressement*, op. cit. págs. 87 ss.
44. Cf. D. Gambetta (ed.), *El sentido de las misiones suicidas*, México, FCE, 2009.
45. A. Downs, *An Economic Theory of Democracy*, N. York, Harper, 1957.
46. B. Keysa, B. Converse, J. Wang, N. Epley, “Reciprocity Is Not Give and Take”, *Psychological Science*, 2008, 19, 12.
47. S. Pinker, *The Blank Slate: The Modern Denial of Human Nature*, Londres, Peguin, 2002.
48. Y que, según parece, también tienen soporte neuronal, cf. J. Lauwereyns, *The anatomy of bias*, Cambridge, Mass., The MIT Press, 2011. Sobre el automatismo de los procesos de decisión en su relación con escenarios sociales e institucionales, cf. Ch. Engel, W. Singer (eds.), *Better Than Conscious?*, Cambridge Mass., The MIT Press, 2008.
49. R. Nickerson, “Confirmation bias: A ubiquitous phenomenon in many guises”, *Review of General Psychology*, 1998, 2. Conviene recordar que la existencia del sesgo no es incompatible con la búsqueda de la verdad. También sabemos —y hay explicaciones evolutivas de ello, faltaría más— que, aunque mentimos más que hablamos a cuenta de nuestras conjeturas, amamos fanáticamente la verdad si hay que maltratar a las ajenas, más exactamente: la gente razona bastante bien a la hora de tasar los argumentos de los otros y bastante mal al defender los propios, H. Mercier, D. Sperber, “Why do humans reason? Arguments for an argumentative theory”, *Behavioral and Brain Sciences*, 2011, 34, 2.
50. Un repaso de los sesgos en G. Gigerenzer, *Decisiones instintivas*, Barcelona, Ariel, 2008.
51. Algunos sesgos se pueden descomponer en otros. Así por ejemplo, la incapacidad para visualizar ignorar el futuro, se sostiene en cuatro disposiciones: a) a atender antes la información vívida que la abstracta; b) a descontar el futuro más allá de lo razonable (pájaro en mano); c) a centrarse en los problemas inmediatos y cercanos, ignorando problemas más graves al fondo, d) a creer que todo se equilibra y acaba compensando, que lleva a pensar en que al final la justicia se impone (“recibirá lo que merece”) o que la naturaleza se repara por sí sola. La “voluntaria ignorancia” ante las crisis económicas o, más seriamente, frente a los problemas ambientales, se ancla en estos mecanismos, cf. G. Markus, “What Should we Worried about?”, *Edge*, 05-04-2013.

52. Lo que contradice resultados importantes de la teoría económica, como el teorema de Coase, según el cual, la disposición a pagar por un bien es la misma que la disposición a cobrar, de modo que, ante un problema de bienes públicos, el único problema es otorgar los derechos de propiedad, y al final, siempre se recala en un precio único de equilibrio, D. Kahneman, J. Knetsch, R. Thaler, “Experimental Tests of the Endowment Effect and the Coase Theorem”, *Journal of Political Economy*, 1990, 98, 6.

53. W. Khanemann, “Willingness to Pay and Willingness to Accept: How Much Can They Differ?”, *The American Economic Review*, 1991, 81, 3.

54. A. Sen, “Rational Fools”, art. cit.

55. La atribución de estados mentales es común a los primates con intensa vida social. Nuestros parientes próximos tienen capacidad para engañar y eso parece exigir una teoría de la mente: “yo quiero que el piense que yo creo que la comida está allí”, cf. D. Cheney, R. Seyfarth, *How monkeys see the world*, Chicago, University of Chicago Press, 1990; C. Allen, M. Bekoff, *Species of Mind*, Cambridge, The MIT Press, 1997, cap. 6; S. Parker, M. McKinney (eds.), *Origins of Intelligence. The Evolution of Cognitive Development in Monkeys, Apes and Humans*. Baltimore: Johns Hopkins, 1999. Para una visión crítica, C. Heyes, “Theory of Mind in Nonhuman primates”, *Behavior and Brain Sciences*, 1998, 21. Su relevancia para la consideración moral, cf. A. Morton, *The Importance of Being Understood: Folk Psychology as Ethics*, Londres, Routledge, 2003.

56. Los pensamientos sobre los pensamientos de otros parecen necesarios también en la “economía animal”, cf. M. Chen, V. Lakshminaryanan, L. Santos, “The evolution of our preferences: Evidence from capuchin monkey trading behaviour”, *Journal of Political Economy*, 2006, 114, 3. De hecho, si esa idea es difícil entender los intercambios, cf. S. Brosnan, F. de Waal, “A concept of value during experimental exchange in brown capuchin monkeys”, *Cebus apella. Folia Primatology*, 2004, 75. Parecen incorporar algo parecido a una idea de valor y, de ahí, se seguiría su rechazo a los intercambios manifiestamente desiguales, S. Brosnan, F. De Waal, “Monkeys reject unequal pay”, *Nature*, 2003, 425.

57. Algo muy dependiente de la existencia de un lenguaje: para que una idea me parezca incorrecta es de estar en condiciones de designarla. Se volverá al final sobre este punto.

58. La formulación es de H. Frankfurt, “Freedom of will and the concept of person” (1971) en G. Watson (edt.), *Free Will*, Oxford, Oxford U. 1982. Cf. también: A. Mele, “Akrasia, Self-Control, and Second-Order Desires”, *Noûs*, 1992, 26, 3. Entre los economistas: T. Schelling, “The intimate contest for self-command”, en T. Schelling, *Choice and Consequences*, Cambridge, Mass. The Harvard U.P. 1984; A. Sen, “Choice, Ordering and Morality” en A. Sen, *Choice, Welfare and Measurement*, 1982, Cambridge, Mass. Harvard University Press, 1997; D. George, “Meta-Preferences: Reconsidering Contemporary Notions of Free Choice”, *International Journal of Social Economics*, 1984, 11, 3. También del gremio: S. Ch. Kokm, *L'homme pluridimensionnel*, París, Albin Michel, 1981. Desde un perspectiva psicológica pero pensando en la economía, G. Ainslie, *Picoeconomics*, Cambridge, Cambridge U.P. 1992. Cf. también J. Elster (edt.), *The multiple self*, Cambridge, Cambridge U.P. 1987. Bueno será advertir que el padre moderno de la criatura sostiene que en el oficio de vivir hay unos fines últimos, unas necesidades volitivas, ante las que “no podemos hacer otra cosa que atenernos de manera incondicional”, entre las que el amor ocupa un lugar paradigmático, H. Frankfurt, *Taking Ourselves Seriously & Getting It Right*, Stanford, Stanford U.P., 2006, pág. 40.

59. O no se decide. No está de más recordar el caso de las gaviotas argénteas que, cuando se encuentran con un huevo de color rojo en su nido (colocado por el experimentador), se vuelven locas. Por un lado, el color rojo activa su agresividad; por otro, el huevo activa su instinto a empollar: intenta atacarlo y a la vez incubarlo. Dos programas independientes impresos en el cerebro que no se traducen en una solución agregada. Cf. D. Eagleman, *Incógnito*, Barcelona, Anagrama, 2013, pág. 175. El ejemplo nos muestra, además, que el problema que nos ocupa es previo a la existencia de preferencias sobre preferencias.

60. La existencia de diversos ordenes de preferencias en su reconstrucción del psicoanálisis por D. Davidson, “Paradoxes of Irrationality”, R. Wollheim y J. Hopkins (eds.), *Philosophical Essays on Freud*, Cambridge, Cambridge U.P., 1982. El problema es asumir que cada uno de ellos se comporta estratégicamente. La idea de “intenciones inconscientes” no es, como tal, inconsistente, pero sí parece serlo la de que el inconsciente tenga un comportamiento estratégico, lo que requiere “pensar sobre lo que no piensa”, esto es, ser consciente. También cabe una formulación, más reciente, a partir de la teoría modular de la mente: cada una de esas unidades especializadas y comparimentadas iría a su aire, con interacciones conflictivas, que se expresaría como inconsistencias, R. Kurzban, *Why Everyone (Else) Is a Hypocrite: Evolution and the Modular Mind*, Princeton, Princeton U.P., 2012.

61. Esta es la reconstrucción en términos de Arrow de I. Steedman y U. Krause, “Goethe’s Faust, Arrow’s possibility theorem and the individual decision taker”, en J. Elster (ed.), *The Multiple Self*, op. cit. De todos modos, cabe pensar que, al final, el agente decide, agrega y actúa: ordena los estados de mundo y dispone de una función de utilidad que, por la definición —moderna— captura todos los pros y los contras. En ese sentido, la propia capacidad de racionalizar, de interpretar acciones que en realidad obedecen a otras razones de las invocadas (ver más abajo nota 80), cumpliría la tarea de “agregación” de preferencias, al menos es algo que se puede colegir de las investigaciones de Michael Gazzaniga sobre el cerebro dividido: todo se integra en un relato único, cf. M. Gazzaniga, *¿Quién manda aquí?*, Barcelona, Paidós, 2012.

62. El resultado inicial es la llamada “paradoja de Easterlin”, según la cual, los aumentos de renta, más allá de cierto umbral, no suponen aumentos de felicidad, R. Easterlin, “Does Economic Growth Improve the Human Lot?” en P. David, M. Reeder, (eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, N. York, Academic Press, Inc, 1974. Eso sí, se trata de un resultado controvertido. Sobre las implicaciones de los resultados empíricos de la felicidad, cf. L. Bruni, P.L. Porta (eds.), *Handbook on the Economics of Happiness*, Cheltenham, UK, Edward Elgar, 2007; E. Posner, C. Sunstein (eds.), *Law and Happiness*, Chicago, The University of Chicago Press, 2010.

63. En realidad lo que se ajusta son las creencias: “la uva está verde”. También sucede con los estudiantes, que encuentran justas notas asignadas al azar: creen que sus respuestas son erradas (o acertadas).

64. Grupo ética, justicia y economía, *Preferencias adaptativas*, Montevideo, Fin de Siglo, 2010; S. Khader, *Adaptive Preferences and Women’s Empowerment*, Oxford, Oxford U.P., 2011.

65. A. Maryanski, J. Turner, *The Social Cage*, Stanford, Stanford U.P., 1992, Para un panorama, A. Maryanski, “Evolutionary Sociology”, *Advances in Human Ecology*, 1998, 7; J. Turner, *On the Origins of Human Emotions*, Stanford, Stanford U.P., 2000. Una excelente discusión del vínculo entre la moral “sociológica” y el sustrato naturalista, evolutivo, en S. Turner, *Explaining The Normative*, Londres, Polity, 2010 (ensayo discutido por varios comentaristas en *Revista Internacional de Sociología*, 2013, 71).

66. La diversidad de criterios normativos está fuera de discusión: L. Furby, “Psychology and justice” en R. Cohen (ed.), *Justice: Views from the social sciences*, N. York, Plenum Press, 1986; R. Shweder, A. Richard, A. Fiske (eds.), *Metatheory in Social Science: Pluralisms and Subjectivities*. Chicago, The University of Chicago Press, 1986; A. Fiske, *Structures of Social Life*, N. York, Free Press, 1991, J. Haidt, “The moral emotions”, en R. Davidson, K. R. Scherer, & H. H. Goldsmith (eds.), *Handbook of affective sciences*, Oxford, Oxford U.P., 2003. Una comparación de las distintas perspectivas en S. Pinker, *Los ángeles que llevamos dentro*, Barcelona, Paidós, 2012, págs. 810 ss.

67. Al menos se pueden concebir escenarios en donde la propia selección natural propicie la coexistencia de individuos con disposiciones diferentes sin que ninguno de ellos llegue a predominar. Un experimento mental lo ejemplificará. En el seno de la población de una especie no agresiva (“palomas”) aparece un gen relacionado con comportamientos agresivos, un “halcón”. Cada vez que un halcón encuentra a una paloma, ésta huye y el otro se queda con el alimento. En esas condiciones los genes de los halcones se extenderán en el seno de la población. Pero hasta un límite. Mientras que los encuentros entre palomas se resuelven con un reparto de comida, cuando se encuentran dos halcones —encuentros que aumentan con el número de halcones se producen enfrentamientos que acaban con la vida de uno de los contendientes y con el otro seriamente dañado. Así las cosas, las palomas empezaran de nuevo a predominar hasta que, de nuevo, con pocos halcones, resultará ventajoso ser halcón y vuelta a empezar. Convivirán “agresivos” y “pacíficos” sin que ninguno de ellos acabe por imponerse en el conjunto de la población, cf. J. Maynard Smith, *Evolution and Theory of Games*, Cambridge: Cambridge U.P., 1982.

68. Hay otro mecanismo por el cual el mercado mina su propia geografía moral: poner precio a las normas. Si se puede comprar o vender todo (las leyes, la confianza), si todo es mercado, desaparece el mercado, que necesita que ciertas cosas estén fuera del juego mercantil. Bienes, como la amistad, la generosidad o la confianza, desaparecen cuando se les pone precio, cf. M. Sandel, *What Money Can’t Buy: The Moral Limits of Markets*, Londres, Allen Lane, 2012.

69. En las sociedades humanas, en todas, están presentes disposiciones cooperativas. Después de todo, la mayor parte de la vida de la especie se ha desenvuelto en escenarios que requerían altos niveles de reciprocidad: defensa, ayuda mutua, coordinación en la caza y recolección. Esos escenarios han ejercido importantes presiones evolutivas que hacían ventajosos en términos adaptativos los comportamientos “morales”. Precisamente porque salen a cuenta para todos pero no resultan interesantes para cada uno, hay que apuntar, para explicarlos, en una dirección distinta del

“cálculo egoísta”. Un repaso de la literatura, eso sí, destacando que hemos sido muy bestias y ahora no tanto, ya con independencia de la biología, en S. Pinker, *Los ángeles que llevamos dentro*, op. cit.

70. El vínculo evolutivo entre emociones y normas están bien documentado entre los homínidos. Simpatía, contagio emocional, culpa, vergüenza, venganza, son algunos de esos soportes emocionales que sirven de cimiento a normas de justicia, de reciprocidad, de minimización del conflicto, a la estabilidad de las relaciones sociales, cf. F. De Wall, *Good Nature*, Cambridge, Mass. Harvard U.P., 1996; *Primates y Filósofos, La evolución de la moral*, Barcelona, Paidós, 2006. En el caso de nuestra especie el salto es radical. El sustrato emocional del comportamiento moral —bien fijado en el “cableado” mental de la especie— es la explicación de una singularidad de los humanos: a pesar de que, en comparación con los monos, los humanos —y los simios, sus parientes más cercanos— presentaban una más baja propensión a la socialidad, unas estructuras sociales más débiles y un mayor grado de autonomía, fueron capaces de organizarse en grupos y sobrevivir en escenarios —en la sabana africana— que reclamaban la organización de amplios grupos. Las presiones selectivas han favorecido en los homínidos conductas que van más allá de la selección de parentesco, hasta extender “el altruismo recíproco a grupos más extensos, hasta otros miembros de la especie”, J. Turner, “The Evolution of Morality”, *Critical Review*, 1997, 11, págs. 218 y 213.

71. A. Damasio, *Descartes’ error: Emotion, reason, and the human brain*, N. York, Putnam, 1994.

72. Algunos neurobiólogos han destacado que “aunque la perspectiva de la toma de decisiones basada en la maximización de utilidad resulta persuasiva, las decisiones humanas rara vez se conforman a ella”. Antes al contrario, sus investigaciones sobre marcadores somáticos muestran que las decisiones “están ancladas en el lado emocional y muy alejadas del constructo del *homo oeconomicus*” A. Bechara, A. Damasio, “The somatic marker hypothesis: A neural theory of economic decision”, *Games and Economic Behavior*, 2005, 52, 2, págs. 336-372.

73. La pregunta la ha formulado con claridad R. Thaler: “La economía neoclásica se sustenta en tres suposiciones: racionalidad, egoísmo y autocontrol (). Sabemos, por supuesto, que esas tres suposiciones son erradas. Los humanos son más estúpidos, más amables y más débiles que el *homo oeconomicus*. ¿Es importante? ¿Es posible hacer economía sin el *homo oeconomicus*? ¿Y si es así, cómo?” R. Thaler, “Doing Economics without *Homo Economicus*”, S. Medema, W. Samuels, (eds.), *Foundations of Research in Economics: How do Economists do Economics*, Edward Elgar, Cheltenham, 1996, pág. 227.

74. La aspiración es “mejorar el poder descriptivo de la teoría económica. El fin de la empresa no debe ser acabar con la economía, si no mejorarla”, R. Thaler, *Ibidem*, pág. 236. Ese punto de vista resulta acorde con tesis realistas defendidas por algunos filósofos de la economía dispuestos a pagar el precio de la precisión y la elegancia formal a cambio de obtener teorías que describan los procesos causales y a las tendencias empíricas, cf. T. Lawson, *Economics and Reality*, Londres, Roudledge, 1997.

75. La hipótesis de la “bounded rationality” cimentada en el reconocimiento del carácter limitado de nuestras capacidades cognitivas, H. Simon, *Models of Bounded Rationality*, 2. Vols., Cambridge U.P., Cambridge., 1982. También las “rules of thumb”, en particular con la “prospect theory” según la cual, los individuos están más atentos a los cambios en la riqueza que a los niveles, D. Kahneman, A. Tversky, D. Kahneman, “Prospect Theory: An Analysis of Decisions Under Risk”, *Econometrica*, 1979, 47, págs. 276-287.

76. S. Bowles, H. Gintis, “*Homo Reciprocans*”, *art. cit.*; S. Bowles, R. Boyd, E. Fehr, H. Gintis, “Explaining Altruistic Behavior in Human”, *Evolution & Human Behavior*, 2003, 24.

77. cf. Fiske, op. cit; Haidt, op. cit.

78. P. Churchland, *Neurocomputational Perspective*, Cambridge, Mass., MIT Press, 1989, pág. 9.

79. P. Churchland, “Folk Psychology and the Explanation of Human Behavior”, en J. Greenwood (edt.), *The Future of Folk Psychology*, Cambridge, Cambridge U.P., 1991. La crítica reduccionista-eliminacionista no tiene una traducción positiva al precisar que teoría económica, y en general, que teoría social, se podría hacer prescindiendo completamente de estados mentales. Por lo general, su argumentación se apoya en analogías y comparaciones de la historia de la ciencia, como la ruptura de la química moderna con respecto a la “química” del flogisto: éste no se explicó o redujo, sino que se descalificó. La nueva teoría no responde a las viejas preguntas, sino que inaugura otras nuevas, imposibles de hacer antes de disponer de ella.

80. En este caso, como casi siempre por lo demás, “genes” es un modo tosco de sintetizar programas impresos en nuestro cerebro que funcionan automáticamente y que son los “responsables” de nuestras acciones, aunque ocasionalmente tengamos la ilusión de gestionar tales acciones. Una versión popularizada de esa idea D. Eagleman, *Incógnito*, Barcelona, Anagrama, 2012. Por estas esquinas rápidamente nos encontramos con el avispero del “libre albedrío”. No está de más decir que cabe una interpretación, bien informada con los resultados de la ciencia, que asume un universo determinista, sin que ello impida contemplar suficiente espacio al ámbito de elecciones personales: H. Walter, *Neurophilosophy of Free Will*, The MIT Press, Cambridge, Mass., 2001. Contra la centralidad de la idea de responsabilidad en el debate moral y político es de interés el trabajo de B. Waller, *Against Moral Responsibility*, Cambridge, Mass., The MIT Press, 2012. Según él, es posible sostener que “el concepto de responsabilidad moral no puede sobrevivir a nuestro mundo naturalista” y que nuestro conocimiento de la naturaleza humana lo descalifica y, a la vez, conservar lo que “realmente nos importa: un mundo natural y no milagroso de libre albedrío, juicio moral, relaciones humanas afectivas y plenas de significado, capacidades creativas y oportunidad para tomar nuestras propias decisiones y ejercer un control efectivo sobre ellas” (pero descartar las nociones de castigo, mérito y recompensa), págs. 1-2.

81. Acciones que obedecen a instrucciones transmitidas exclusivamente a un hemisferio cerebral pero que el otro hemisferio, interrogado y en aislamiento, se ve en la necesidad de explicar, de inventarse un relato mediante creencias y deseos, cf. M. Gazzaniga, *The Social Brain*, N. York, Basic Books, 1985, pág. 80. Aun más, la propia generación de esa “teoría económica popular” se podría explicar desde la propia biología. Sería un modo sofisticado de relato psicológico, de construir “teorías” acerca de por qué han ocurrido comportamientos que responden a otras causas.

82. Y aquí cabría incluir también la naciente investigación rotulada como “neuroeconomía” que, por así decir, salta por encima de los supuestos intencionales para intentar hacer pie en la neurociencia, para “medir”, por ejemplo, la utilidad prescindiendo de lo que digan los sujetos, a partir de neuroimágenes, cf. P. Glimcher, *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*, Cambridge, Mass., The MIT Press, 2003. Sobre ese debate, defendiendo que la economía no tiene ningún problema, porque, en realidad, no incorpora supuestos psicológicos sobre el cerebro, F. Gul, W. Pesendorfer, “The Case for Mindless Economics”, en A. Caplin, A. Schotter (ed.), *The Foundations of Positive and Normative Economics*, Oxford, Oxford U.P., 2010. Una valoración matizada de la disciplina en el número monográfico (preparado por G. Bonanno, Ch. List, B. Tungodden y P. Vallentyne) de *Economics and Philosophy*, 2008, 24, 3.

83. Cierto es que hay una teoría económica (a la que se ha bautizado como enfoque reproductivo) que viene de la línea Ricardo-Marx-Sraffa, en la que se hace uso de modelos con microfundamentos sin apelar inevitablemente a estados mentales, al modo como sucede con las tablas *input-output*. En ese caso, los precios se “explican” no tanto por la interacción de agentes maximizadores, compradores y vendedores, como por las condiciones técnicas de producción. Pero incluso ahí, en algún momento, asoman los agentes, a través, por ejemplo, de supuestos sobre una igual tasa de beneficios que responde a la competencia.

84. Entrevista, en M. Szenberg, *Grandes economistas de hoy*, Barcelona, Debate, 1994, págs. 52-53.

85. Y eso, entre otras cosas, nos permite sobrevivir. En lugar de ensayar con nosotros y arriesgarnos a morir, ensayamos con nuestras construcciones intelectuales, D. Dennet, “The Role of Language in Intelligence”, J. Khalfa (ed.), *What is Intelligence?*, Cambridge, Cambridge U.P., 1994.

86. Más en general, cualquier teoría que pretende dar cuenta de ciertos sucesos en algún sentido interesantes (E) apelando a ciertas condiciones causales (C), se puede entender como una herramienta crítica: la situación (la realidad) C* resulta criticable en la medida que está alejada de C, la que permite obtener el resultado que se juzga interesante E. Y también como una prescripción normativa (si se quiere que suceda E, constrúyanse las condiciones C) de la que se siguen principios prácticos: debemos modificar C* en dirección a C. En ese sentido, la teoría del equilibrio general, muchas veces descalificada por su irrealidad empírica, también se puede entender como una teoría que nos dice que el mercado real es incapaz de asegurar ciertos objetivos, en la medida que está en E*, alejado de las condiciones ideales (y que siempre lo estará), dado que muchas de esas condiciones E son de imposible cumplimiento. Así las cosas, el mercado perfecto se revelaría como una utopía comparable al socialismo perfecto. A eso se añade el resultado, bien conocido, de que estar cerca de E, de la situación deseable, pero no estar en ella, puede ser peor que estar lejos, R. Lipsey, K. Lancaster “The General Theory of Second Best” (1957), en R. Kuenne, ed., *Readings in Social Welfare: Theory and Policy*, Blackwell, Oxford, 2000. Una combinación de fármacos, en el que falta uno, puede ser letal.

87. Cf. M. Gladwell, *La inteligencia intuitiva: ¿Por qué sabemos la verdad en dos segundos?*, Madrid, Taurus, 2005.

88. Para autores como S. J. Gould y N. Chomsky no es una adaptación específica, sino el subproducto de otras adaptaciones. Tesis criticada por S. Pinker y P. Bloom en su clásico trabajo “Natural Language and Natural Selection”, *Behavioral and Brain Sciences*, 1990, 13, 4.

89. J. Hurford, *The Origins of Meaning Vol.1 of Language in the Light of Evolution*, Oxford, Oxford U.P., 2007, págs. 22-70.

90. La recursividad resulta fundamental, en su relación con la teoría de la mente, resulta en esto fundamental: ese yo se que tu sabes que yo pienso, cf., Michael C. Corballis, *The Recursive Mind: The Origins of Human Language, Thought, an Civilization*, Princeton, Princeton, 2011.

91. E. Stein, *Whitout Good Reason*, Oxford, Clarendon Press, 1996.

Fecha recepción: 21/04/2013

Fecha aceptación: 21/10/2013