

El egoísmo psicológico *

ELLIOT SOBER

Hans Reichenbach and Vilas Research Professor
Universidad de Wisconsin-Madison

El egoísmo psicológico es una teoría sobre la motivación que afirma que nuestros deseos últimos son autocentrados. Las críticas contra el egoísmo psicológico se pueden dividir en tres categorías: *a)* se dice que no es una auténtica teoría; *b)* que es una teoría refutada por la observación de la conducta; *c)* que se debería rechazar en favor de una teoría alternativa según la cual los seres humanos tienen deseos

últimos tanto egoístas como altruistas. Se analizan estos tres tipos de crítica y se concluye que la situación del debate entre egoísmo y pluralismo motivacional es de tablas. Situación que puede encontrar alguna salida a partir de consideraciones evolucionistas. El egoísmo no merece ser considerado como hipótesis por defecto. Aunque sea en un grado pequeño, el peso de la evidencia favorece al pluralismo.

El egoísmo psicológico es una teoría sobre la motivación que afirma que todos nuestros deseos últimos están dirigidos hacia nosotros mismos. Siempre que queremos hacer bien (o mal) a otros tenemos estos deseos orientados hacia los demás sólo de manera instrumental; nos preocupamos por los otros solamente porque pensamos que su bienestar tendrá ramificaciones o consecuencias sobre el nuestro. En su forma habitual el egoísmo es una propuesta descriptiva, no normativa. Pretende caracterizar aquello que motiva de hecho a los seres humanos. La teoría no nos dice si es bueno o malo que las personas tengan tales motivos.

El egoísmo ha ejercido una influencia poderosa en las ciencias sociales, y ha calado profundamente en la opinión común. En particular, los economistas habitualmente consideran que los humanos son seres movidos por su «auto-interés racional» y, con ello, excluyen cualquier otra preocupación irreductible por el bienestar de otros. La opinión común habitual sostiene normalmente que unas personas ayudan a las otras sólo porque esto les hace sentirse bien consigo mismas o porque con ello buscan la aprobación de terceras partes.

Incluso para los actos más impresionantes de sacrificio personal resulta fácil inventar explicaciones egoístas. El soldado que metido en una trinchera se arroja sobre una granada para salvar la vida de sus camaradas es una constante en la literatura sobre el egoísmo. ¿Cómo es posible que ese acto sea producto de su propio interés, si el soldado sabe que le llevará a una muerte segura? El egoísta puede responder que el soldado se da cuenta, por un instante,

* Traducción de J. Francisco Álvarez.

que prefiere morir antes que sufrir los sentimientos de culpa que le acosarían si se salvara y dejase perecer a sus compañeros. El soldado prefiere morir, y no tener ningún tipo de sensación posterior, antes que vivir y sufrir los tormentos de la culpa. Esta réplica puede parecer forzada, pero resta analizar qué bases tenemos para considerarla falsa.

Las críticas que se han dirigido contra el egoísmo psicológico se pueden dividir en tres categorías. En primer lugar, la que afirma que no es una auténtica teoría. Una segunda crítica consiste en decir que es una teoría refutada por lo que observamos en la conducta humana. En tercer lugar, aunque el egoísmo sea una teoría consistente con lo que observamos, nos encontramos con la opinión de que hay otras consideraciones extraevidenciales que sugieren que debería rechazarse a favor de una teoría alternativa, el pluralismo de motivos, según la cual los seres humanos tienen deseos últimos tanto egoístas como altruistas. En lo que sigue vamos a analizar estos tres tipos de crítica, pero, previamente, es conveniente formular la teoría con mayor precisión.

1. *Aclaraciones sobre el egoísmo*

Cuando con el egoísmo se afirma que todos nuestros deseos últimos están dirigidos hacia el propio yo, ¿qué significa «último» y «autocentrados»?

Hay ciertas cosas que queremos por ellas mismas, mientras que hay cosas que deseamos solamente porque pensamos que con ellas conseguiremos algunas otras. La conocida relación medios/fines, que vincula un deseo a otro, permite también que los deseos se encadenen entre sí —Sara puede querer conducir su coche porque quiere ir a la panadería, puede querer ir a la panadería porque quiere comprar pan, etc.—. La relación crucial que se precisa definir es la siguiente:

S quiere *m* solamente como un medio para adquirir *e*, si y solamente si *S* quiere *m*, *S* quiere *e* y *S* quiere *m* solamente porque cree que conseguir *m* le ayudará a conseguir *e*.

Un deseo último es simplemente un deseo que alguien tiene por razones que están más allá de la capacidad que ese deseo tenga para contribuir instrumentalmente al logro de alguna otra cosa. Consideremos el dolor. La razón más obvia por la que la gente quiere evitar el dolor es simplemente porque les disgusta experimentarlo. Evitar el dolor es uno de nuestros objetivos finales. De todas formas, muchas personas se dan cuenta de que tener dolor reduce su capacidad de concentración, de manera que a veces se toman una aspirina debido en parte a que quieren eliminar una fuente de distracción. Esto muestra que las cosas que queremos como fines en sí mismos también pueden quererse por razones instrumentales.

Cuando el egoísmo psicológico pretende explicar por qué una persona ayudó a otra no es suficiente mostrar que *una* de las razones para ayudarle fue la

del propio beneficio; esto es plenamente consistente con que los individuos tengan otras razones, puramente altruistas, para ayudar. Simétricamente, para refutar el egoísmo no es preciso poner ejemplos de ayuda en las cuales representen un papel sólo aquellos motivos exclusivamente orientados hacia otros. Si las personas ayudan a veces por razones últimas tanto egoístas como altruistas entonces el egoísmo psicológico es falso.

Tanto el egoísmo como el altruismo requieren la distinción entre deseos dirigidos hacia uno mismo y deseos orientados hacia otros. Esta distinción se debe entender en términos del contenido proposicional del deseo. Si Adán quiere la manzana, estamos ante una forma elíptica de decir que Adán quiere que ocurra que *él tenga la manzana*. Este deseo es puramente autorreferencial, puesto que su contenido proposicional menciona a Adán y a ningún otro agente. Por el contrario, cuando Eva quiere que *Adán tenga la manzana* este deseo es puramente dirigido hacia otros, su contenido proposicional menciona a otra persona, Adán, pero no a Eva. El egoísmo afirma que todos nuestros deseos últimos son autocentrados; el altruismo plantea que algunos de esos deseos están dirigidos hacia otros. El hecho de que Eva tenga un deseo dirigido hacia otro no es suficiente para rechazar el egoísmo, debemos plantearnos *por qué* Eva quiere que Adán tenga la manzana.

Una versión especial del egoísmo es el hedonismo psicológico. El hedonista dice que los únicos deseos últimos que tienen las personas son obtener placer y evitar el dolor. A veces se critica al hedonismo porque sostiene que el placer en una sensación de tipo único, que son la misma cosa el placer que conseguimos por saborear un melocotón y el que se obtiene al ver que prosperan en algo aquellos a quienes amamos (Lafollette, 1988). En todo caso, esta crítica no se aplica al hedonismo tal como lo hemos descrito. El hecho destacable de esta teoría es que afirma que las personas son solipsistas motivacionales; en último término lo único que les preocupa son los estados de su propia consciencia. Los egoístas no tienen por qué ser hedonistas. Si las personas se preocupan por su propia supervivencia como un fin en sí mismo puede que sean egoístas, pero no hedonistas.

Hay deseos que ni están totalmente dirigidos hacia uno mismo ni tampoco están exclusivamente orientados hacia los otros. Que Phyllis quiera ser famosa significa que quiere que otros sepan quién es. El contenido proposicional de este deseo presupone una relación entre el yo y los otros. Si Phyllis busca la fama exclusivamente porque piensa que le resultaría placentero o beneficioso, puede que sea un egoísta (dependiendo de cuáles sean sus deseos últimos). Pero, ¿qué decir si quiere ser famoso como un fin en sí mismo? No hay ninguna razón por la cual debemos repartir esta posibilidad entre el egoísmo o el altruismo. De manera que podemos reconocer al *relacionismo* como una posibilidad diferente de las otras dos. Construido de esta manera, el egoísmo evita así la dificultad de tener que explicar por qué la teoría es compatible con la existencia de algunos deseos últimos relacionales, pero no con otros (Kavka, 1986).

Una cuarta posibilidad consiste en suponer deseos que no hacen mención ni de uno mismo ni de los otros. El deseo de mantener y defender algún principio moral general cae dentro de esta categoría. Cuando un utilitarista desea el mayor bien para el mayor número, el deseo es impersonal; el deseo cubre a todos los seres sentientes, presumiblemente incluye también al propio individuo que tiene ese deseo, pero el contenido del deseo no señala ni al propio individuo ni a ningún otro ser específico. Por esta razón, sugiero que no es ni egoísta ni altruista. Lo mismo que era cierto con respecto a los deseos relacionales, el defensor del egoísmo psicológico puede aceptar que haya deseos referidos a los principios morales generales que no son dirigidos hacia sí mismo, la cuestión es si tenemos estos deseos de manera instrumental o como fines en sí mismos.

Es evidente que los individuos actúan sobre la base de sus propios deseos y tratando de lograr la satisfacción de esos deseos, pero también es obvio que de estas dos verdades no se deduce el egoísmo caracterizado tal como he sugerido. Que Joe actúe sobre la base de los deseos de Joe, no sobre la base de los de Jim, nos dice cuáles deseos (y de quién) están actuando, pero no nos dice nada sobre si los deseos últimos que alberga Joe en su cabeza son exclusivamente autodirigidos. Que Joe quiera que sus deseos se satisfagan significa exclusivamente que quiere que se cumpla su contenido proposicional; el deseo de Joe de que llueva mañana se satisface si y solamente si mañana llueve (Stampe, 1994). Si llueve, el deseo se satisface, sepa o no Joe que ha ocurrido así. Querer que el deseo de uno se satisfaga no es lo mismo que querer conseguir el sentimiento de satisfacción que a veces acompaña al deseo satisfecho.

A veces se critica al egoísmo porque atribuye demasiado cálculo a los actos espontáneos de ayuda. Las personas que ayudan en situaciones de emergencia con frecuencia declaran haberlo hecho «sin pensar» (Clark y Word, 1974). En cualquier caso, es difícil aceptar literalmente esas declaraciones, cuando los actos suponen una serie precisa de acciones complicadas realizadas adecuadamente para obtener un fin bien claro. El socorrista que rescata a un nadador en dificultades puede considerarse adecuadamente como alguien que tiene un objetivo y que selecciona las acciones que permiten lograrlo. Que no se le considere inmerso en un cálculo bien ponderado y autoconsciente no muestra que no se haya producido algún razonamiento en término de medios y fines. En todo caso, las acciones que ocurren de hecho sin la mediación de creencias y deseos caen fuera del ámbito del egoísmo lo mismo que fuera del espacio del altruismo. Las personas extienden sus piernas cuando se les golpea sus rodillas con un martillo, pero esto no refuta ninguna teoría.

Otra crítica relacionada con la anterior es la que sostiene que el egoísmo supone que las personas son más racionales que lo que realmente son. Vale la pena recordar que el egoísmo es simplemente una afirmación sobre los deseos últimos que tienen las personas. Como tal no dice nada sobre cómo

los individuos deciden qué hacer sobre la base de sus creencias y deseos. Los teóricos que asumen que el egoísmo es verdadero también aceptan con frecuencia que las personas son calculadores racionales; sin embargo, a las teorías no se les acusa por un principio de culpabilidad por asociación. No es mayor la vinculación del supuesto de racionalidad con el egoísmo psicológico que la vinculación que tiene dicho supuesto con el pluralismo de motivos.

Si el egoísmo sostiene que todos los deseos últimos están orientados hacia sí mismo, ¿qué podremos decir de alguien cuyo fin último sea su propia destrucción? Y, si el altruismo sostiene que algunos de nuestros deseos últimos están dirigidos hacia los otros, ¿cómo podemos abordar el caso de Yago, cuyo objetivo último consiste en la destrucción de Otelo? Parece muy extraño calificar de egoísta a una persona deprimida que llega al suicidio, lo mismo que decir que Yago es un altruista. Hace falta añadir a las dos teorías una idea de lo que es el bien (o aparentemente bueno). Los egoístas buscan su propio beneficio, los altruistas quieren que otros estén bien. Aunque estos añadidos a las correspondientes teorías las sitúan más cerca del uso ordinario de los términos «egoísmo» y «altruismo», no afectan de hecho a la cuestión substantiva de determinar cuál de las dos teorías es la correcta. La clave del problema es determinar si todos los deseos últimos son deseos dirigidos hacia uno mismo.

Puede que algunos lectores consideren que el problema es fácil. Sencillamente, los individuos pueden estar atentos a lo que ocurre dentro de sus propias mentes y determinar introspectivamente cuáles son sus motivos últimos. Quizá los partidarios del egoísmo tengan razón sobre ellos mismos y posiblemente también la tengan los defensores del pluralismo motivacional por lo que se refiere a ellos; ambos sólo se equivocan cuando generalizan más allá de sus propios casos. Un supuesto implícito en las investigaciones filosóficas y psicológicas de este asunto es la aceptación de que las personas son básicamente iguales. Si el egoísmo es falso, es falso para casi todos (exceptuando quizás a los sociópatas). Y, si esto es cierto, lo es porque caracteriza un rasgo básico de la naturaleza humana. Sin embargo, que los estudios previos en psicología y filosofía hayan ignorado con frecuencia la posibilidad de variación individual no es ninguna razón seria para convertirlos en la base de nuestra comprensión del problema. ¿Por qué no decir que los partidarios del egoísmo conocen sus propios corazones y que lo mismo ocurre con los defensores del altruismo?

La clave está en que no hay ninguna razón independiente para pensar que el testimonio de la introspección sea confiable en este caso. Por lo que respecta a otros aspectos de la mente, la introspección es confusa e incompleta; nadie ha demostrado por qué la mente sin embargo tendría que ser un libro abierto en relación con esta cuestión de los motivos últimos.

El problema, en el caso de que se pueda resolver, debería ser abordado de otra manera.

2. *¿Es el egoísmo empíricamente contrastable?*

Una típica objeción filosófica al egoísmo consiste en decir que no es una hipótesis contrastable. Como sugiere el ejemplo del soldado metido en la trinchera, parece que el egoísmo puede ajustarse a cualquier clase de conducta. Ya sean los individuos desagradables o amables con los otros, la teoría puede explicar por qué. Esta afirmación de la flexibilidad del egoísmo se enlaza fácilmente con el criterio popperiano relativo a las exigencias que debe cumplir un enunciado para considerarlo científico, llevándonos a la conclusión de que el egoísmo no es de ninguna manera una auténtica teoría científica. A pesar de todas las apariencias, el egoísmo es una teoría empíricamente vacía.

Sin embargo, el argumento resulta inadecuado en un doble sentido. En primer lugar, se trata de la aceptación confiada de que ninguna observación jamás puede refutar el egoísmo. Que la teoría se pueda acomodar al caso del soldado en la trinchera y a otras conductas propuestas por los filósofos es manifiestamente insuficiente para justificar esa afirmación general. De hecho, sucede que el trabajo experimental en psicología social, a propósito del altruismo y del egoísmo, muestra que la evidencia observacional pertinente cubre un campo más amplio que la existencia de casos de conducta solidaria o de ayuda (Batron, 1991; Schroeder y otros, 1995).

Además, la observación de Duhem de que las teorías solamente son contrastables conjuntamente con un conjunto de supuestos que conforman su trasfondo contextual nos lleva a abandonar esta crítica de la no contrastabilidad. Si dos teorías con un mismo trasfondo contextual hacen las mismas predicciones, bien pudiera ocurrir que hicieran predicciones diferentes si tuvieran otro marco contextual. ¿Cómo podremos saber qué nuevos contextos teóricos se desarrollarán y que permitirán poner a prueba el egoísmo? La acusación de no contrastabilidad presupone que tenemos un control omnisciente del futuro de la ciencia.

El segundo defecto de este argumento consiste en que no se percata de que la acusación de incontrastabilidad es un arma de doble filo. El argumento se propone como una razón para rechazar el egoísmo. Entonces, ¿qué aceptar como explicación positiva de la conducta? Posiblemente puede suponerse que el pluralismo de motivos sea una alternativa razonable. En todo caso, ésta no puede ser la conclusión del argumento. Si el egoísmo no es contrastable, tampoco lo será el pluralismo de motivos. Si el egoísmo resulta muy flexible en su capacidad para adaptarse a las diversas observaciones, el pluralismo motivacional lo es incluso más. Después de todo, el pluralismo incorpora todas las variables que utiliza el egoísmo, e incluso alguna más. Las dos teorías se relacionan entre sí de la misma manera que lo hacen « $y = f(x)$ » e « $y = g(x, w)$ ».

El egoísmo parece que no es contrastable porque es un *-ismo*. No suministra explicaciones específicas de las conductas, sino que simplemente señala el tipo de explicación que deberían tener todas las conductas.

Ésta es la razón por la que se puede seguir manteniendo el egoísmo aunque se consideren insatisfactorias las explicaciones concretas que se apoyan en él. ¿Por qué Jorge donó todo su dinero para obras benéficas? Un defensor del egoísmo puede plantear que Jorge así lo hizo porque quería impresionar a otros para, con ello, mejorar sus relaciones comerciales. Supongamos que alguien sabe que Jorge hizo la donación en forma anónima. Con esto se refutaría aquella concreta explicación que se apoyaba en el egoísmo, pero no resulta nada difícil inventarse otra. Por ejemplo, Jorge hizo la donación porque le hacía sentirse bien consigo mismo y porque sabía que de no hacerlo así experimentarían un remordimiento culpable. El patrón de explicación ahora utilizado es típico —el hedonismo es la posición a la que normalmente se retiran los egoístas—. Si los beneficios externos no resultan suficientes como explicación, se invocan beneficios psicológicos internos.

El egoísmo es una afirmación sobre un *tipo* de explicación y, por ello, es diferente de las *explicaciones específicas* del tipo requerido. Este modelo aparece en muchos debates sobre los *-ismos*. Pensemos en el adaptacionismo en biología evolucionista. Los adaptacionistas enfatizan la importancia de la selección natural a la hora de explicar los rasgos observados de los organismos. Debido a que el *-ismo* como tal no suministra ninguna explicación específica de ningún rasgo, queda abierta la posibilidad para que un biólogo continúe siendo adaptacionista incluso después de que una determinada explicación adaptacionista se considere inadecuada. ¿Por qué evolucionaron las alas en los insectos? La hipótesis de que las alas evolucionaron como una adaptación para volar se pone en duda por el hecho de que existen alas muy pequeñas totalmente inservibles para volar; aunque el 5 por 100 de un ojo puede seguir funcionando como sensor de luz, el 5 por 100 de un ala no hace nada para facilitar que un organismo levante el vuelo. Sin embargo, se encuentran alas en algunas especies de insectos no voladores; funcionan como termorreguladores. Este hecho sugiere una hipótesis adaptacionista alternativa —las alas de los insectos comenzaron a evolucionar debido a que inicialmente facilitaban la termorregulación y luego continuaron funcionando porque facilitaban el vuelo—. El adaptacionista, si ve que se ataca a esta última hipótesis, puede articular y presentar otra tercera posición. Sin embargo, no es conveniente rechazar el adaptacionismo porque presente este tipo de flexibilidad; el *-ismo* alternativo, el pluralismo evolucionista, afirma que la selección natural es una causa entre otras de la evolución. El adaptacionismo es flexible, pero el pluralismo lo es incluso más.

3. *El guijarro de Butler*

Como ya he señalado, aunque el hedonismo sea una versión especial del egoísmo, con frecuencia los defensores del egoísmo utilizan las explicaciones hedonistas cuando resultan inadecuadas las no hedonistas. Si Jorge no donó el dinero para obras benéficas con el fin de hacer buenos contactos comerciales, quizá lo hizo por la agradable satisfacción que le produjo la donación. Por esta razón los argumentos que intentan refutar el hedonismo ocupan un lugar muy particular en la controversia que nos preocupa. Aunque no sea suficiente refutar el hedonismo para rechazar el egoísmo, sin duda sería una importante contribución para esa tarea más amplia. Muchos filósofos han pensado que Joseph Butler (1692-1752) refutó el hedonismo de una vez para siempre (Broad, 1965; Feinberg, 1984; Nagel, 1970) en el fragmento siguiente:

Que todos los apetitos y pasiones están dirigidos hacia *las cosas externas*, diferentes del *placer* que brota de ellas, resulta manifiesto de todo lo anterior; pues bien pudiera no darse este placer, si no fuera por esa previa vinculación agradable entre el objeto y la pasión: no se podría disfrutar más o deleitarse más con una cosa que con otra, al alimentarse más que al tragarse una piedra, si no hubiera una afección o apetito mayor por una cosa que por otra (Butler, 1726, p. 227).

Llamaré a este argumento el argumento del guijarro de Butler. Aunque Butler no dice explícitamente en ese fragmento que el hedonismo sea falso, nos permite construir un argumento que conduce a dicha conclusión:

1. Las personas a veces experimentan placer.
2. Cuando las personas experimentan placer es debido a que tienen deseo de alguna cosa externa y de que ese deseo se satisfizo.

El hedonismo es falso.

No propongo que ataquemos la primera premisa. Sin embargo, me parece que la segunda premisa es falsa y que la conclusión no se sigue de las premisas.

La segunda premisa abarca demasiado; aunque algunos placeres sean resultado de la satisfacción de un deseo, otros no lo son (Broad, 1965, p. 66). Podemos disfrutar del aroma de las violetas sin habernos formado previamente el deseo de oler una flor o algo dulce. Puesto que los deseos son actitudes proposicionales, la formación de un deseo es un logro cognitivo. Además, el placer y el dolor a veces están mediados cognitivamente pero en otras ocasiones no lo están. Obsérvese que este defecto del argumento se puede recomponer, pues a Butler no le hace falta decir que la satisfacción del deseo sea una forma y la única de acceso al placer.

En el paso de las premisas a la conclusión es dónde realmente falla el argumento. Considérese la cadena causal que a partir de un *deseo* (el deseo

de comida, por ejemplo) lleva a una *acción* (el comer) y de ahí a un *resultado* —placer—. Debido a que el placer remite a un deseo previamente existente, sería falso considerar que el placer resultante causase el deseo (en el supuesto de que la causa debe preceder al efecto). Sin embargo, esto no nos resuelve cómo se relacionan los dos deseos —el deseo de comida y el deseo de placer—. En particular, deja completamente abierta la cuestión de qué fue lo que causó el deseo de comida. El hedonismo dice que las personas desean comida *porque* quieren placer (y piensan que la comida les traerá ese placer). El guijarro de Butler concluye que esta cadena causal es falsa, pero no lo hace con buenas razones. El error decisivo en el argumento procede de confundir dos asuntos bien diferentes —el *placer* que resulta de la satisfacción de un deseo y el *deseo de placer*—. Aunque cuando se dé el placer presupongamos que el agente deseaba algo además del placer, de ello no se sigue nada referente a la conexión entre el *deseo de placer* y el deseo de alguna otra cosa (Sober, 1992; Stewart, 1992; Sober y Wilson, 1998). El hedonismo no niega que la gente desee cosas externas; más bien lo que la teoría trata de explicar es por qué ocurre eso.

Resulta curioso que este argumento se haya interpretado de manera tan general como refutación del hedonismo. Al final del sermón en el que aparece el párrafo del argumento del guijarro, Butler dice lo siguiente: «Aunque la virtud y la rectitud moral consisten efectivamente en la atracción por lo que es justo y su búsqueda como tal; admitiremos incluso que cuando nos encontramos en horas bajas en las que no podemos justificarnos para con nosotros mismos éste o cualquier otro propósito, todavía estamos convencidos de que será por nuestra felicidad o al menos que no será contrario a ella» (Butler, 1726, p. 240). Si volvemos al lenguaje del argumento, vemos que Butler está haciendo una afirmación sobre el contenido de «particulares apetitos y pasiones». Leído con rigor, el argumento dice simplemente que si la gente desea placer sus deseos no caen bajo esa rúbrica; el argumento no dice que las personas no desean nunca placer ni tampoco dice que el deseo de placer no sea nunca último. ¿Fallaría Butler a la hora de refutar el hedonismo con el argumento del guijarro porque no era ése su propósito?

4. La «paradoja» del hedonismo y su «irracionalidad»

Los individuos que se preocupan exclusivamente por la obtención de placer o felicidad inevitablemente fracasan en lograr lo que pretenden. Son como agentes de bolsa que solamente piensan en comprar a la baja y vender al alza. Quienes tienen un fin a la vista pero nunca se preocupan de los medios que deberían utilizar para conseguirlo, con toda seguridad fracasan en obtener lo que quieren. Esto ha producido que muchos filósofos consideren que el placer y la felicidad solamente se pueden conseguir como subproductos vinculados a actividades específicas. También han sugerido que ese rasgo del placer y la felicidad constituye una paradoja para el hedonismo —la palabra

«paradoja» pretende sugerir que encontraremos aquí ciertos fallos en el hedonismo como teoría psicológica (Butler, 1726; Feinberg, 1984).

La réplica obvia a tal tipo de crítica consiste en decir que no hay nada en el hedonismo que diga que las personas deban ser monomaniacas. El hedonismo dice que los individuos tienen como únicos objetivos *últimos* el obtener placer y evitar el dolor; pero no dice que conseguir placer y evitar dolor sean los únicos objetivos (últimos o próximos) que siempre tengan las personas. El hedonista reflexiona sobre qué actividades son más aptas para producir placer y prevenir el dolor y decide qué hacer sobre tal base (Sidgwick, 1907). Más aún, si el hedonista monomaniaco fracasa siempre en conseguir lo que quiere, ¿qué se sigue de ello? Incluso si esto implicase que las personas no *deberían* ser hedonistas, con ello no se demostraría que no lo sean *de hecho*. Debemos recordar que el hedonismo es una teoría descriptiva y no una teoría normativa.

La misma distinción normativo/descriptivo hace falta también para evaluar la afirmación de que el egoísmo es irracional. Nagel (1970) defiende esta idea al proponer que cuando los egoístas consideran sus propios intereses en la deliberación y, a la vez, no tienen en cuenta los de los demás, no se fijan en que no hay propiedad alguna que tengan ellos, y no los demás, que pueda justificar esa asimetría. Para evaluar si los egoístas son irracionales, debemos decidir si la racionalidad ha de entenderse «instrumental» o «substantivamente». La racionalidad instrumental significa capacidad para elegir medios eficientes que permitan alcanzar cualesquiera fines que uno pueda tener. La noción substantiva no significa que se hayan asegurado los medios eficientes, sino que los fines sean dignos de valorar o que al menos sean moralmente inobjectables (Gibbard, 1990). Los asesinos en serie eficientes puede que sean instrumentalmente racionales, pero no son racionales substantivamente. Con independencia de cuál sea la noción que capte mejor lo que signifique la palabra «racional», sigue sucediendo que esta línea argumental no puede mostrar que las personas de hecho tengan (o sean capaces de tener) motivos últimos de tipo altruista. Si la racionalidad significa racionalidad instrumental, la racionalidad no entraña el altruismo (o su posibilidad); y si la racionalidad significa racionalidad substantiva, incluso aunque la racionalidad entrañe el altruismo, falta por demostrar que las personas son en realidad substantivamente racionales. Quizá *debamos* ser racionales y puede que también *debamos* ser altruistas. Esto no demuestra que el egoísmo sea falso como tesis descriptiva.

5. *La máquina de la experiencia*

En la película de ciencia-ficción *Desafío total* las gentes de siglos venideros utilizan su tecnología informática para irse de «vacaciones virtuales». En vez de salir realmente de vacaciones, se enganchan a un ordenador que les ofrece una simulación muy convincente de unas auténticas vacaciones. La película

sugiere de manera muy plausible que esas gentes del futuro con frecuencia podrán elegir tomar «vacaciones» de esa forma, en particular si los viajes reales a lugares exóticos resultan muy caros y peligrosos mientras que las «vacaciones virtuales» son baratas y totalmente convincentes desde el punto de vista de la experiencia.

Robert Nozick escribió *Anarquía, estado y utopía* mucho antes del estreno de *Desafío total*. En el libro aparece la idea de una «máquina de la experiencia» para elaborar un argumento que parece demostrar que el hedonismo es falso (Nozick, 1974, pp. 42-45). La máquina de Nozick puede programarse para ofrecer simulaciones altamente convincentes de cualquier experiencia de la vida real que uno quisiera elegir. Supongamos que se nos ofrece la oportunidad de engancharnos a esa máquina de la experiencia para el resto de nuestra vida. La máquina podría estar programada para que instantáneamente nos olvidásemos de que hemos elegido engancharnos a ella y para que luego nos ofrezca cualquier tipo de sucesión de experiencias que pudiéramos considerar máximamente placenteras y mínimamente dolorosas. Por supuesto, nuestras creencias sobre el tipo de vida que estaríamos llevando serán falsas. Si elegimos engancharnos a la máquina de la experiencia, viviremos nuestra vida atados a una mesa de laboratorio con tubos y electrodos pegados a nuestro cuerpo. Nunca *haríamos* otra cosa; pero, en cualquier caso, el nivel de placer que experimentaríamos, gracias a la máquina, sería extraordinario.

Si se nos ofreciera la oportunidad para engancharnos a una tal máquina de la experiencia por el resto de nuestra vida, ¿qué haríamos? La primera reacción podría ser de duda sobre si la máquina cumplirá su cometido tan bien como nos lo prometen; es cierto que por ahora no hay ninguna máquina en el mercado que pueda desarrollar lo que se dice que esta máquina es capaz de hacer, y esto seguirá siendo así al menos en un futuro previsible. De todas formas, por continuar con la argumentación, tratemos de obviar esta duda. Imaginemos que nos ofrecen la oportunidad de engancharnos a la máquina y que funciona tan bien como la hemos descrito. Me parece que muchas personas, quizá incluyéndole a usted, rechazarían la oportunidad de engancharse.

Puede parecer que este rasgo de las personas refute el hedonismo. Aparentemente muchas personas prefieren tener una vida real antes que una simulada, incluso aunque la vida real les produzca menos placer y más dolor que la vida que podrían llevar si se conectaran a la máquina. Parece que nos preocupamos irreductiblemente sobre cómo nos relacionamos con el mundo exterior a nuestras mentes; es falso que la única cosa que nos preocupa como fines en sí mismos sean los estados de conciencia placenteros.

¿Puede el hedonismo explicar por qué muchas personas rechazarían la oferta de conectarse a la máquina? Para ver si esto es posible necesitamos poner en correspondencia la secuencia de sucesos que conformarían nuestras vidas si eligiésemos conectarnos a la máquina de la experiencia y la secuencia que se daría en el caso de decidir no conectarnos. En ambos casos el proceso

que éstos esperan que se les va a provocar *después* de que hayan decidido hacerlo, la afirmación que debería sustentar es la de que $a + b < n_1 + n_2$. Puesto que b es mucho mayor que n_2 , la desigualdad se cumplirá solamente si a es muchísimo menor que n_1 . Es decir, el hedonista parece que está obligado a defender que los individuos rechazan la opción de conectarse debido a que es *gigantesca* la cantidad de dolor que experimentarían entre el momento de haber decidido conectarse y el instante en que efectivamente se conectan a la máquina —tan enorme, que empequeñece el placer que pudieran experimentar después de haberse conectado.

No parece muy aceptable esta línea de pensamiento. El período de tiempo que transcurre entre el momento de la decisión y la conexión efectiva puede ser muy breve en comparación con el gran número de años que podría pasar unido a la máquina disfrutando de un conjunto de experiencias máximamente placenteras. Acepto como un hecho que quienes decidan conectarse a la máquina experimentarán una gran tristeza y ansiedad durante el breve intervalo de tiempo que transcurre entre la decisión y la conexión real. Sin embargo, la idea de que esta experiencia negativa inunde y supere todo el placer posterior no parece creíble.

William Talbott me sugirió otro segundo experimento mental que puede servirnos para ver mejor las razones anteriores. Supongamos que nos ofrecen un millón de pesetas si pasamos durante diez segundos por determinada experiencia. La experiencia consiste en creer que hemos decidido pasar el resto de nuestra vida conectados a la máquina de la experiencia. Después de los diez segundos de haber sufrido dicha experiencia, volveremos a nuestra vida normal y nos daremos cuenta de que todo ha sido una «pesadilla», y recibiremos el dinero prometido. Supongo que muchas personas elegirían los diez segundos descritos debido a que ganarían el millón de pesetas. Esto nos muestra que el hedonismo está equivocado si afirma que la experiencia de creer que estaríamos conectados a la máquina de la experiencia por el resto de nuestra vida es tan horrible que nadie elegiría nunca una vida que la incluyera.

El hedonista no ha sido capaz de explicar por qué muchos elegirían una vida normal antes que otra de conexión a la máquina de la experiencia. La razón está en que el cálculo hedonista parece conducir inevitablemente a la conclusión de que $a + b > n_1 + n_2$. ¿Significa esto que el hedonista tiene que aceptar que le hemos refutado? Me parece que le queda una salida. Con independencia de la cantidad de placer y de dolor que acumulen los sujetos *después* de que hayan decidido qué hacer, debemos considerar el nivel de placer y de dolor que surge del proceso mismo de deliberación. El hedonista puede sostener que *decidir* conectarse a la máquina resulta tan aversivo que los individuos casi siempre adoptan la otra opción. Cuando deliberan sobre las alternativas sucede que se sienten muy mal cuando piensan en la vida que llevarían si se conectasen a la máquina, y se sienten mucho mejor cuando consideran la vida que llevarían en el mundo real si rechazaran conectarse.

La *idea* de una vida enlazada a la máquina resulta dolorosa, incluso cuando tal tipo de vida pudiera estar plena de placer; la *idea* de una vida real es placentera, aunque con frecuencia en la vida real tengamos dolor. Esta explicación hedonista de por qué las personas rechazan conectarse explota la distinción trazada por Schlick (1939) entre el placer que acompaña a la idea de un cierto estado de cosas y la idea de un estado placentero.

Para comprender mejor lo que se supone con dicha sugerencia, analicemos con mayor detalle lo que pasa por las mentes de las personas cuando deliberan. Se dan cuenta de que si se conectan abandonan los proyectos y las actividades que anhelan; conectarse recuerda al suicidio, por lo que se refiere a la separación radical que se produce con el mundo real. La diferencia radica en que el suicidio supone poner fin a la consciencia, mientras que la máquina de la experiencia conduce (literalmente) a placeres escapistas. El hedonismo no se ve amenazado en sus propios principios cuando afirma que muchos sentirían una gran satisfacción con la idea misma de conectarse pero considera aborrecible la tentación de hacerlo. Quienes rechazan la oportunidad de conectarse lo hacen porque repelen la idea de un escape narcisista y encuentran placer en la idea de elegir una vida real.

Una virtud de esta explicación hedonista se halla en que explica los resultados que se dan en los dos experimentos mentales descritos. Explica por qué los individuos con frecuencia rechazan conectarse a la máquina de la experiencia para el resto de sus vidas, y también explica por qué normalmente a quienes se le ofrece el millón suelen *aceptar* tener durante diez segundos la experiencia de creer que han decidido conectarse a la máquina para el resto de sus vidas. En ambos casos la deliberación se ve orientada por el placer y el dolor que acompaña a ciertos pensamientos *durante el proceso de deliberación*, y no tanto por las creencias sobre las acciones que producirán *placer futuro*.

El problema de la máquina de la experiencia recuerda el problema ya discutido del soldado en la trinchera. ¿Cómo puede el hedonismo explicar ese acto de autosacrificio suicida, si el soldado cree que no tendrá ninguna experiencia después de la muerte? El hedonista puede sugerir que hay un beneficio autocentrado que se acumula o disfruta *antes* de haber realizado el acto de autosacrificio. Para nada se violenta el hedonismo al mantener que el soldado decide sacrificar su vida debido a que tal decisión es menos dolorosa que la decisión de dejar que mueran sus compañeros. El problema del autosacrificio suicida y el problema que plantea la máquina de la experiencia pueden afrontarse de la misma manera.

6. *La carga de la prueba*

A veces los filósofos defienden que una idea de sentido común debe considerarse inocente hasta que no se demuestre su culpabilidad. Es decir, si se plantea que alguna proposición de sentido común es verdadera, y no se da ningún

argumento que la justifique o la recuse, entonces lo adecuado es mantener nuestra creencia en esa proposición. Dicho de otra manera, defienden que la carga de la prueba cae del lado de quienes atacan el sentido común.

Esta actitud general aparece con frecuencia en la discusión sobre el egoísmo y el altruismo. Suele decirse que la hipótesis del egoísmo es contraria al sentido común. Se mantiene que la imagen de sentido común sobre la motivación humana es pluralista —las personas se preocupan de sí mismas, pero también se preocupan por los demás como fines en sí y no sólo como medios—. La conclusión que se obtiene es que si la argumentación filosófica y científica, a favor y en contra del egoísmo, no resulta decisiva, entonces deberíamos rechazar el egoísmo y continuar aceptando el pluralismo.

Una objeción a este desenlace es que no resulta nada obvio que «el sentido común» se encuentre más cercano al pluralismo de motivos que en la proximidad del egoísmo. ¿Qué es el sentido común? ¿No es lo que normalmente creen las personas? Si es así, se puede defender que el egoísmo tiene recorrido un largo camino, parece ser una perspectiva que acepta gran número de personas. Los filósofos deberían tener cuidado en no confundir el sentido común con lo que ellos mismos consideran obvio. En lo que se me alcanza, ningún tipo de estudio empírico ha logrado determinar si una teoría pluralista de la motivación es más popular que el egoísmo psicológico.

Sin fijarme en lo que la gente cree normalmente sobre el egoísmo psicológico y el pluralismo de motivos, rechazo la idea de que la conformidad con el sentido común sea lo concluyente para este debate. Esa adecuación no tiene lugar ni en física ni en biología, y no veo razón alguna para que debiera tener un papel cuando las cuestiones que se debaten son de carácter filosófico o psicológico. De hecho, se puede plantear que nuestras intuiciones en este dominio son particularmente propensas al error. Los individuos tienen una imagen de sus propios motivos y de los de los otros. Si determinados tipos de auto-decepción —en la consideración de los propios motivos o de los de los otros— resultasen ventajosos, entonces la evolución podría haber incorporado estas falsedades en el conjunto de proposiciones «obvias» que llamamos sentido común. Una filosofía orientada por una perspectiva evolucionista no tiene por qué aceptar a primera vista el sentido común.

7. Parsimonia

En todo lo que llevamos discutido no hemos refutado el hedonismo con argumentos filosóficos ni por la conducta observada; si lo hemos hecho bien, tampoco hemos refutado el egoísmo. Esto no significa que el egoísmo sea verdadero; después de todo, tampoco hemos refutado el pluralismo de motivos. A la vista de tal *impasse* vale la pena observar que los científicos sociales con frecuencia asumen de manera implícita que si una conducta *puede* explicarse en términos egoístas, entonces *debería* explicarse de esta manera. El hecho de que no haya

argumentos directos en favor de esta posición no parece ser algo relevante. Sin embargo, ¿por qué habría de ser el egoísmo la hipótesis por defecto que deberíamos aceptar como verdadera a menos que nos viéramos forzados a abandonarla?

Una posible respuesta a considerar es que el egoísmo es más parsimonioso —postula solamente un tipo de motivos últimos, mientras que el pluralismo postula dos (Hume, 1751; Batson, 1991)—. Incluso si asumimos que la parsimonia apunta a algo que va más allá de una simple diferencia estética entre dos teorías, y que es una razón para encontrar que algunas teorías son más plausibles que otras, sigue habiendo un defecto en esta defensa del egoísmo. El problema es que el egoísmo es *menos* parsimonioso que el pluralismo cuando consideramos cuántas creencias causales postula cada una de las teorías. Cuando Sally quiere que Otto haga las cosas bien, el defensor del egoísmo contabiliza esto como un deseo instrumental, mientras que el proponente del pluralismo motivacional puede defender que Sally tiene como un fin en sí mismo este deseo dirigido hacia otro. Pero debemos fijarnos en que la explicación egoísta atribuye a Sally una creencia causal —*que espera recibir un beneficio de que le vaya bien a Otto*—. El pluralismo de motivos no se compromete con que Sally tenga esa creencia. Un egoísta tiene una lista de deseos últimos menor que la sostenida por el pluralista, pero el egoísta tiene una lista mayor de creencias causales. Por esta razón, no está nada claro por qué el egoísmo psicológico debiera considerarse como la teoría más parsimoniosa (Sober y Wilson, 1998).

8. *Un tratamiento evolucionista*

Los motivos psicológicos son mecanismos próximos en el sentido en que se utiliza el término en biología evolucionista. Cuando un girasol gira hacia el sol debe darse algún mecanismo en el interior del girasol que cause que las cosas ocurran de esa manera. Por tanto, si el fototropismo es una adaptación que evolucionó porque dotaba al organismo de ciertos beneficios, entonces también ha debido evolucionar algún mecanismo próximo que cause esa conducta. De igual manera, si ciertas formas de conducta de ayuda entre los seres humanos son adaptaciones evolutivas, entonces los motivos que causan esas conductas en los seres humanos individuales también han debido evolucionar. Quizá una perspectiva general sobre la evolución de los mecanismos próximos podrá arrojar alguna luz sobre el problema específico de si es más probable que haya evolucionado el egoísmo o el pluralismo motivacional.

Seguir este tratamiento evolutivo no presupone que cada detalle de la conducta humana, o cada acto de ayuda, pueda explicarse completamente con la hipótesis de la evolución por selección natural. Sin duda hay muchos aspectos de la conducta, y muchos ejemplos de ayuda, para los cuales la selección natural no es una explicación pertinente. Sin embargo, quiero considerar un deter-

minado hecho sobre la conducta humana y mi afirmación es que la selección es pertinente para explicarlo. El aspecto interesante que pretendo señalar es que los progenitores humanos cuidan de sus hijos, que la cantidad promedio de cuidado dado a sus descendientes por parte de los seres humanos es manifiestamente mayor que el que dan otros progenitores en otras muchas especies. Supondremos que la selección natural ofrece al menos parte de la explicación de por qué el cuidado de los progenitores evoluciona en nuestro linaje evolutivo. Con esto no negamos que entre los progenitores humanos no haya variabilidad, que algunos cuidan más de sus hijos que otros y que algunos incluso abusan y matan a su propia descendencia. Otro hecho llamativo sobre la variación individual es que las madres, en promedio, invierten más tiempo y esfuerzo en ese cuidado que el empleado por los padres. Quizá también existan explicaciones evolutivas para esas diferencias individuales; la cuestión que quiero plantear aquí, sin embargo, no adopta ningún supuesto sobre si éste es el caso.

Repasemos algunos principios generales que orientan cómo se pueden predecir cuáles serán los mecanismos próximos que evolucionarán para producir una conducta particular. Cambiaré de asunto y me fijaré en unos ejemplos relativos a unos organismos hipotéticos sin especiales capacidades cognitivas cuyo problema consiste en seleccionar del entorno determinados materiales para alimentarse. Algunas partículas que flotan en el medio líquido en el que vive el organismo contienen proteínas, otras tienen veneno. El organismo ha evolucionado hacia una conducta particular —tiende a comerse las proteínas y a evitar el veneno—. ¿Qué mecanismo próximo ha podido evolucionar para permitirles hacer las cosas así?

En primer lugar veamos el conjunto de soluciones posibles que necesitamos tener en cuenta. El diseño de solución más obvio para este problema consiste en que el organismo tenga un detector que permita distinguir las proteínas del veneno. Capta un bocado que flota, coloca la partícula en su detector y asocia el resultado del detector a una conducta; el organismo se traga el bocado o lo expulsa. A esta solución la llamaré solución *directa* para el problema de diseño; el organismo necesita discriminar entre proteínas y veneno, y esta solución alcanza ese fin utilizando un detector que revela precisamente ese contraste entre las propiedades.

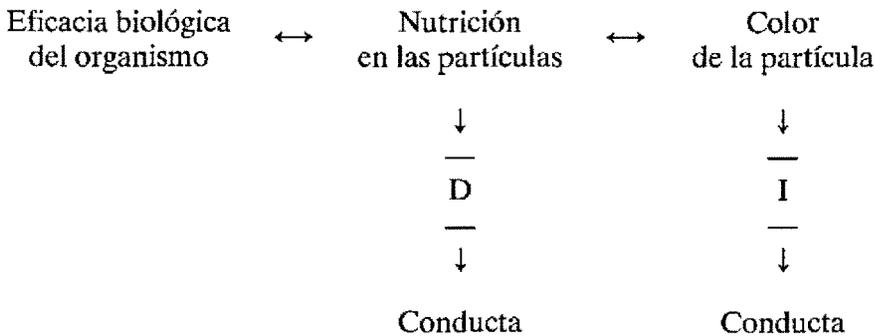
No resulta difícil imaginar otras soluciones menos directas para este problema de diseño. Supongamos que en el entorno del organismo las proteínas tienden a ser rojas y que el veneno tiende a ser verde. Si las cosas son así, el organismo podría utilizar un detector de colores para conseguir la discriminación que precisa. Esta solución diseñada es *indirecta*, el organismo necesita distinguir las proteínas del veneno, y para conseguirlo discrimina, entre otras, dos propiedades que están correlacionadas con el objetivo a diferenciar. En general, puede que haya muchas soluciones de diseño indirectas que pudiera explotar el organismo; hay tantas soluciones indirectas como correlaciones pue-

da haber entre la distinción proteína/veneno y otras propiedades que se den en el entorno. Por último, podríamos añadir a nuestra lista la idea de que pueden darse soluciones pluralistas a un determinado problema de diseño. Además de las soluciones monistas como la de tener un detector de proteínas o la de tener un detector de colores, un organismo podría exhibir un detector de proteínas y un detector de colores.

Dada la multiplicidad de posibilidades, ¿cómo podemos predecir cuál de ellas será la que evolucionará? Para este asunto resultan pertinentes tres principios: *disponibilidad*, *fiabilidad* y *eficiencia* (Sober, 1993; Sober y Wilson, 1998).

La selección natural actúa solamente en el ámbito de la variación que existe previamente entre los ancestros. Podría ser una buena cosa que el organismo tuviera un detector de proteínas, pero si ese artificio no estuvo nunca presente como variante hereditaria, la selección natural no puede conseguir causalmente que evolucione ese rasgo. De manera que el primer tipo de información que nos gustaría conseguir se refiere a qué mecanismos próximos estaban *disponibles* para poder ser heredados.

Supongamos, para seguir con el argumento, que se pudiera disponer hereditariamente tanto del detector de proteínas como del detector de colores. ¿Cuál de ellos es más probable que evolucione? Necesitamos fijarnos en el asunto de la *fiabilidad*. ¿Qué mecanismo consigue hacer de manera más fiable el trabajo de indicar qué partículas del entorno son buenas como alimento? Sin mayor información no se puede avanzar mucho. Un detector de colores puede tener cualquier grado de fiabilidad y lo mismo puede decirse de un detector de proteínas. No hay ninguna razón *a priori* por la cual la estrategia directa sea más o menos fiable que la estrategia indirecta. Sin embargo, hay una circunstancia especial en la que difieren. Se puede ilustrar con el siguiente diagrama:



Las dobles flechas indican correlación; los beneficios nutricionales se correlacionan con la eficacia biológica (*fitness*) del organismo, y que una partícula sea roja en lugar de verde se correlaciona con su contenido nutritivo. En el diagrama no hay flecha que vaya desde la eficacia biológica al color, salvo

una que pasa a través de la nutrición. Esto quiere decir que la eficacia biológica del organismo se correlaciona con el color de la partícula que se come. No hay ninguna razón *a priori* por la cual el color debiera ser pertinente para la eficacia biológica, salvo por el hecho de que indica el contenido nutricional. Por ejemplo, si comer partículas rojas atrae a los predadores más que si se comen las verdes el color podría tener dos tipos de pertinencias para la eficacia biológica o valor adaptativo. Sin embargo, si la nutrición «filtra» la eficacia biológica procedente del color de la manera indicada, podemos establecer el siguiente principio sobre la fiabilidad del mecanismo directo D y la del mecanismo indirecto I:

(D/I) Si la nutrición y el color no están perfectamente correlacionados, y si D detecta la nutrición al menos tan bien como I detecta el color, entonces D será más fiable que I.

Éste es el principio de asimetría directo/indirecto. Las soluciones directas para un problema de diseño no son siempre las más fiables, pero son más fiables en estas circunstancias.

También a partir del anterior diagrama puede derivarse otro segundo principio relativo a la fiabilidad. De la misma manera que los científicos hacen mejor su trabajo de discriminación entre hipótesis si disponen de una evidencia mayor, así también sucede que los organismos ejecutan discriminaciones más fiables si en lugar de una sola fuente de información tienen dos fuentes de información sobre qué les conviene comer.

(DMU) Si la nutrición y el color no están perfectamente correlacionadas, y si D e I son ambos detectores fiables de la nutrición, aunque falibles, ocurre que D e I serán más fiables si hacen el trabajo conjuntamente que si cada uno trabajase por su cuenta.

Este es el principio Dos es Mejor que Uno. Necesita un supuesto, que los dos mecanismos no interfieran entre sí cuando se presentan conjuntamente en un mismo organismo; que funcionen de manera claramente independiente.

La asimetría D/I y el Principio DMU pertenecen al ámbito de la fiabilidad. Veamos ahora una tercera consideración pertinente para predecir cuál de los mecanismos próximos será el que evolucione, es decir, la *eficiencia*. Incluso aunque se disponga de un detector de nutrientes y de un detector de colores, e incluso aunque el detector de nutrientes sea más fiable, de esto no se sigue que la selección natural favorecerá al detector de nutrientes. Podría suceder que la construcción y el mantenimiento del funcionamiento de un detector de nutrientes requiriese mayor cantidad de energía que la necesaria para un detector de colores. Los organismos funcionan con energía en medida no menor que los automóviles. La eficiencia es tan pertinente para el rasgo general de eficacia biológica como lo es la fiabilidad.

Teniendo en cuenta las tres consideraciones anteriores, volvamos ahora al problema de predecir cuál de los mecanismos motivacionales, capaces de producir el cuidado por parte de los progenitores, es probable que haya evo-

lucionado en el linaje evolutivo que conduce a los seres humanos. Los tres mecanismos motivacionales que debemos analizar se corresponden con tres reglas diferentes (hedonista, altruista y pluralista) adecuadas para seleccionar una conducta a partir de lo que uno cree:

(HED) Dar el cuidado como progenitores si y solamente si al hacerlo se maximiza el placer y se minimiza el dolor.

(ALT) Dar el cuidado como progenitores si y solamente si al hacerlo así se consigue mejorar el bienestar del propio hijo.

(PLUR) Dar el cuidado como progenitores si y solamente si al hacerlo así o bien se maximiza el placer y se minimiza el dolor, o bien se consigue mejorar el bienestar del propio hijo.

(ALT) es una solución relativamente directa y (HED) una relativamente indirecta al problema de diseño para conseguir que un organismo se preocupe por su descendencia. Lo mismo que un organismo puede alimentarse detectando el color, también es posible en principio que un organismo hedonista se construya de manera tal que pueda ofrecer cuidado a su descendencia; lo que hace falta es que el organismo esté constituido de forma tal que el dar ese cuidado sea lo que maximice su placer y minimice su dolor (o que al menos el organismo crea que así es la cosa).

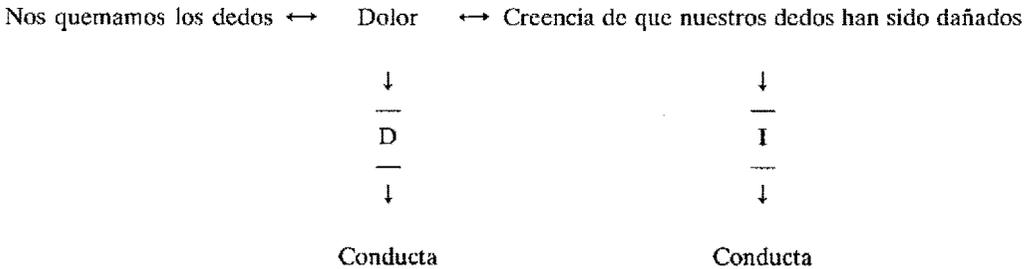
Consideremos el grado de fiabilidad que tendrían esos tres mecanismos en una determinada situación. Supongamos que un progenitor sabe que su descendiente está en peligro. Imaginemos que un vecino nos dice que nuestro hijo se ha caído por entre los hielos en un lago helado. (HED) y (ALT) realizan su tarea de la manera siguiente:

El niño necesita ayuda ↔ Los padres creen que el niño necesita ayuda ↔ Los padres sienten ansiedad y temor



Los padres altruistas se verán impulsados a la acción porque creerán que su hijo necesita ayuda. Los padres hedonistas no lo harán así; más bien lo que les moverá a la acción serán los sentimientos de ansiedad y temor que les causa la noticia o la creencia de padre de que tales sentimientos negativos le acompañarán siempre a menos que se mejore la situación del niño. Resulta claro a partir del diagrama que se aplica en este caso el principio de asimetría (D/I). En estas circunstancias concretas, (ALT) será más fiable que (HED). Además, por el principio (DMU), (PLUR) lo hará mejor que ambos. En este ejemplo el hedonismo aparece como tercero en la competición, al menos en tanto que nos preocupe la fiabilidad.

La importancia de este ejemplo radica en que los sentimientos que tienen los padres son *creencias mediadas*. La única razón por la que los padres sienten ansiedad y temor es porque los padres *creen* que su hijo tiene problemas. Así ocurre también en muchas de las situaciones en las que se baraja el egoísmo o el hedonismo como explicación, pero no es cierto en todas. Por ejemplo, consideremos la siguiente situación de daño corporal en la cual el dolor es un efecto directo y la creencia es un efecto relativamente indirecto de ese daño:



Ahora ocurre que el hedonismo es una solución directa al problema de diseño. No tendría sentido construir un organismo de manera que no respondiera al dolor y que retirase sus dedos de la llama solamente después de haberse formado una creencia sobre el daño corporal. En esta situación, *la creencia está mediada por el dolor* y el principio de asimetría (D/I) explica por qué tiene pleno sentido una atención hedonista sobre el dolor. Sin embargo, el mismo principio señala que resulta erróneo considerar al hedonismo una solución diseñada para cuando el dolor venga mediado por la creencia, que es precisamente lo que ocurre con tanta frecuencia en el contexto del cuidado y atención realizada por parte de los progenitores.

Si el hedonismo es menos fiable que el puro altruismo o que el pluralismo motivacional, ¿cómo podremos comparar los tres mecanismos cuando consideramos los problemas evolutivos de la disponibilidad y la eficiencia? Con respecto a la disponibilidad, me gustaría hacer la siguiente afirmación: *Si el hedonismo se encontraba como una opción disponible entre nuestros ancestros, como una posible solución a diseñar, lo mismo puede decirse del altruismo*. La razón está en que los dos mecanismos motivacionales solamente difieren en algo muy menor. Los dos precisan de una psicología de creencias/descos. Tanto los padres altruistas como los hedonistas quieren que a sus hijos les vaya bien; la única diferencia está en que el hedonista tiene su contenido proposicional como un deseo instrumental mientras que el altruista lo tiene como deseo último. Si el altruismo y el pluralismo no han evolucionado, esto no ha sido porque no estuvieran disponibles para que la selección actuara sobre ellos.

¿Qué decir a propósito de la eficiencia? ¿Es necesario un mayor número de calorías para construir y mantener un organismo altruista o pluralista que

para construir y mantener un organismo hedonista? No veo por qué. Lo que exige energía es la construcción de un sistema físico que sea capaz de incorporar una psicología de creencias/deseos. Sin embargo, es difícil ver por qué tener un deseo último en lugar de dos pudiera suponer alguna diferencia energética; tampoco es fácil ver por qué tener el deseo último de que nuestro hijo se encuentre bien exija mayor número de calorías que tener el deseo último de evitar el dolor y obtener placer. En apariencia, las personas con más creencias no necesitan comer más que quienes tienen un número menor. Lo mismo parece que se aplica al asunto de cuántos, o cuáles, deseos últimos se tienen.

En resumen, el hedonismo es un mecanismo menos fiable que el altruismo puro o que el pluralismo en tanto que mecanismo para desarrollar el cuidado paterno. Por lo que respecta a los temas de disponibilidad y de eficiencia, no nos parece que haya diferencias entre estos tres mecanismos motivacionales. Esto sugiere que la selección natural es más probable que nos haya hecho pluralistas motivacionales antes que hedonistas.

Desde un punto de vista evolutivo, el hedonismo es un mecanismo motivacional muy tosco. Lo que importa en el proceso de selección natural es la capacidad de un organismo para sobrevivir y tener éxito reproductivo. El éxito reproductivo no significa simplemente la producción de descendencia, sino la supervivencia de aquellos supervivientes que alcancen la edad reproductiva. De manera que lo importante es la supervivencia del propio cuerpo y de los cuerpos de los hijos de uno. El hedonismo, por otra parte, dice que los organismos se cuidan en último término de los estados de su propia consciencia y solamente de eso. ¿Por qué debería la selección natural haber conducido a organismos que se preocupan de algo que es periférico a la eficacia biológica (*fitness*), en lugar de fijar su atención en el precio? Si los organismos fueran incapaces de conceptualizar proposiciones sobre sus propios cuerpos y los cuerpos de sus descendientes, ésa pudiera ser una razón. Después de todo, para un organismo puede tener sentido explotar la estrategia indirecta de decidir qué comer sobre la base del color en lugar de apoyarse en el valor nutricional, si los organismos no tienen ningún acceso directo al contenido nutritivo. Pero si un organismo es lo suficientemente inteligente como para formarse representaciones sobre sí mismo y su descendencia, esta justificación de la estrategia indirecta no resulta plausible. Que hayamos evolucionado desde nuestros ancestros que eran menos complejos, cognitivamente hablando, hace que no sea sorprendente que evitar el dolor y conseguir placer sean dos de nuestros fines últimos. Pero que los seres humanos sean capaces de formarse representaciones con muy diversos contenidos proposicionales sugiere que la evolución ha añadido más piezas a la lista de lo que nos preocupa como fines en sí mismos.

9. Comentarios finales

He defendido que no han tenido éxito los intentos filosóficos y psicológicos del pasado para intentar resolver el debate entre egoísmo y pluralismo motivacional. Sería realmente sorprendente que esta disputa sobre un asunto aparentemente empírico pudiera resolverse por argumentos *a priori*. Desgraciadamente, las observaciones que los individuos hacen en su vida cotidiana y las que hacen los científicos en el laboratorio tampoco han resultado decisivas. Aunque algunas versiones simples del egoísmo se refutan por lo que observamos, otras versiones del egoísmo se pueden construir de manera que parece que se ajustan a las observaciones disponibles. Quizá experimentos y observaciones más complejas de la conducta pudieran responder la cuestión. Pero, por ahora, la situación en filosofía y en psicología es de tablas.

¿Podieran las consideraciones evolutivas romper el *impasse*? La argumentación desarrollada en la sección previa pretende plantear que es menos probable que haya evolucionado un conjunto de motivos puramente egoístas que uno que incluya descos últimos tanto egoístas como altruistas. No pretendemos sugerir que este argumento *pruebe* que las personas sean pluralistas motivacionales; hay mucho por aprender sobre la mente y cómo ha evolucionado, y no tenemos garantía de que posteriores detalles no cambien el cuadro que he tratado de dibujar. Sin embargo, me parece que el argumento es suficiente para demostrar que el egoísmo no merece ser considerado como la hipótesis por defecto que deberíamos adoptar en la medida en que resulta consistente con lo que observamos. En mi opinión, aunque sea en un grado pequeño, el peso de la evidencia favorece al pluralismo.

BIBLIOGRAFÍA

- BATSON, C. D.: *The Altruism Question: Toward a Social-Psychological Answer* (Hillsdale, N. J.: Lawrence Erlbaum Associates, 1991).
- BROAD, C. D.: *Five Types of Ethical Theory* (Totowa, N. J.: Littlefield, Adams, 1965).
- BUTLER, J.: *Fifteen Sermons upon Human Nature* (1726), reprinted in L. A. Selby-Bigge (ed.), *British Moralists: Being Selections From Writers Principally of the Eighteenth Century*, vol. 1 (New York: Dover Books, 1965), pp. 180-241.
- CLARK, R. D., y WORD, L. E.: «Where is the apathetic bystander? Situational characteristics of the emergency», *Journal of Personality and Social Psychology*, 29 (1974), pp. 279-287.
- FEINBERG, J.: «Psychological egoism», en S. Cahn, P. Kitcher, and G. Sher (eds.), *Reason at Work* (San Diego, C. A.: Harcourt Brace and Jovanovich, 1984), pp. 25-35.
- GIBBARD, A.: *Wise Choices, Apt Feelings* (Cambridge: Harvard University Press, 1990).
- HUME, D.: «On self love», en *An Enquiry Concerning the Principles of Morals* (1751) (Indianápolis: Hackett, 1970).
- KAVKA, G.: *Hobbesian Moral and Political Theory* (Princeton, N. J.: Princeton University Press, 1986).

- LAFOLLETTE, H.: «The truth in psychological egoism», en J. Feinberg (ed.), *Reason and Responsibility*, 7.^a ed. (Belmont, C. A.: Wadsworth, 1988), pp. 500-507.
- NAGEL, T.: *The Possibility of Altruism* (Oxford: Oxford University Press, 1970).
- NOZICK, R.: *Anarchy, State, and Utopia* (New York: Basic Books, 1974).
- SCHLICK, M.: *Problems of Ethics* (New York: Prentice Hall, 1939).
- SCHROEDER, D.; PENNER, L.; DOVIDIO, J., y PILIAVIN, J.: *The Psychology of Helping and Altruism* (New York: McGraw-Hill, 1995).
- SIDGWICK, H.: *The Methods of Ethics* (1907), 7.^a ed. (London: Macmillan, 1922).
- SOBER, E.: «Hedonism and Butler's stone», *Ethics*, 103 (1992), pp. 97-103.
- SOBER, E.: «Did evolution make us psychological egoists?», en *From a Biological Point of View: Essays in Evolutionary Philosophy* (New York: Cambridge University Press, 1994), pp. 8-27.
- SOBER, E., y WILSON, D. S.: *Unto Others-the Evolution and Psychology of Unselfish Behavior* (Cambridge: Harvard University Press, 1998).
- STAMPE, D.: «Desire», en S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*, (Cambridge, M. A.: Basil Blackwell, 1994), pp. 244-250.
- STEWART, R. M.: «Butler's argument against psychological hedonism», *Canadian Journal of Philosophy*, 22 (1992), pp. 211-221.