

LA JUSTIFICACIÓN BIOLÓGICA DE LA MORAL: MODELOS Y PERSPECTIVAS*

ALEJANDRO ROSAS
UNIVERSIDAD NACIONAL DE COLOMBIA

Resumen:

Contra la tesis de que una explicación biológica de la moralidad es irrelevante para su justificación o conduce a su socavamiento (Th. Nagel), arguyo que el concepto de función biológica abre un camino que lleva desde la biología hasta el “espacio lógico de las razones (prácticas)”. Luego reconstruyo dos modelos filosóficos para la recepción de la sociobiología, que coinciden en poner al altruismo psicológico en el núcleo de la moral. Explico sus diferencias y critico la restricción de la relevancia de la sociobiología a mecanismos de tipo emotivo, arguyendo que la moral basada en un altruismo condicionado sugerida por la sociobiología bien puede fundarse en mecanismos cognitivos.

Palabras claves: Biología; moral; Nagel; McGinn; Mackie.

Abstract: *The Biological Justification of Morality: Models and Perspectives.*

I argue against the claim that a biological explanation of morality is either irrelevant to justification or implies a debunking explanation (Th. Nagel), and that the concept of biological function opens a way that leads from biology to the “logical space of (practical) reasons”. I then reconstruct two models for the philosophical reception of the sociobiology of altruism, both of which view altruism at the center of morality. I explain their differences and criticize their limitation of sociobiology’s relevance to emotive mechanisms, arguing that a morality based on conditional altruism, as suggested by sociobiology, could well rest on cognitive mechanisms.

Key words: Biology; morality; Nagel; McGinn; Mackie.

1. La filosofía moral en perspectiva biológica

Avances recientes en la biología evolucionista presentan a la selección natural como un proceso que debe entenderse desde la perspectiva de los genes como unidades de selección (George C. Williams, William D. Hamilton y Robert Trivers). Se trata de un desarrollo científico con un gran impacto potencial en la filosofía moral y social. Al menos esa es la pretensión que le ha valido celebridad filosófica a E.O. Wilson y a su primer libro sobre la evolución del comportamiento social: *Sociobiology. The New Synthesis* (1975), pero también a las publicaciones de Richard Dawkins, entre las que se destaca *El Gen Egoísta*

* La investigación para este ensayo fue posible gracias al apoyo académico e institucional de la Universidad Nacional de Colombia y del profesor Ludwig Siep de la Universidad de Münster, Alemania; y al apoyo financiero de la Fundación *Alexander von Humboldt*. Agradezco también al árbitro anónimo de *Ideas y Valores* por sus sugerencias, que contribuyeron a mejorar el ensayo.

(1976). Este es el espíritu que anima afirmaciones como ésta, del biólogo Richard Alexander:

Creo que todo concepto relevante a la socialidad humana (como racionalidad, conciencia, culpa, autoconciencia, altruismo y egoísmo) cambia su sentido –o lo precisa– al aplicarle los nuevos refinamientos de la teoría evolutiva (Alexander 1987:3)

En su *The Biology of Moral Systems*, Alexander sostiene que la explicación biológica de la moralidad, basada en la tesis de que el gen es la unidad de la selección, destapará una verdad ingrata acerca de nosotros mismos, a saber, que la moralidad es un mecanismo deceptivo para fines egoístas. Conviene examinar con cuidado esta tesis, que es presentada como inevitable para quienes aceptan la idea de que la evolución debe verse desde la perspectiva del gen. Teóricamente, las consecuencias filosóficas de una explicación biológica de la moral podrían clasificarse del modo siguiente: 1) socavar los fundamentos de la moral y destaparla como una ilusión de algún tipo; 2) teniendo en cuenta que hay diversas concepciones de la moral –algunas más exigentes que otras– la explicación biológica podría dar fundamento a las más modestas y socavar las más exigentes como ilusorias e ilegítimas; 3) podría justificar incluso las más exigentes; o, por último, 4) podría ser irrelevante para la justificación y dejar todo como está. Nótese que todas ellas suponen que una explicación biológica de la moralidad que apela a la selección natural es posible. Esta suposición es difícil de evitar en nuestro tiempo, y de ello da testimonio el hecho de que aun los filósofos más conservadores procuran acomodar sus teorías a esa posibilidad. Thomas Nagel es un buen ejemplo.

2. Thomas Nagel y la autonomía de la Razón

Quizás no sea imposible, en el marco de la concepción evolutiva del ser humano, defender a la razón como un poder autónomo y asiento de la actitud moral. Thomas Nagel lo intenta en su libro *The Last Word* (1997). Su tratamiento seduce por su simplicidad, pero esconde, visto de cerca, una posición inestable.

Nagel está aparentemente dispuesto a aceptar una explicación evolutiva de la racionalidad, pero le niega un carácter justificatorio con dos argumentos. Señala primero, que si la explicación se ofreciese también como justificación sucumbiría a la objeción de circularidad, pues toda explicación tiene que apoyarse en la razón, y no podría justificarla sin suponer su previa validez (*Id.*, 134): “La hipótesis evolutiva es aceptable sólo si la razón no necesita apoyarse [justificarse o legitimarse, A.R.] en ella” (*Id.* 136). Esta es una objeción débil, pues puede hacersele a todo intento de justificación de la razón: tiene que

apoyarse en la razón y es por tanto circular. En segundo lugar, Nagel sostiene que la explicación evolutiva termina por minar la confianza en la validez de la razón, es decir, es incompatible con el realismo que él defiende.

El argumento para esta segunda tesis es el siguiente. En opinión de Nagel, Darwin proveyó a la cultura secular de una "manera de eliminar propósito, significado y diseño como rasgos fundamentales del mundo" (*Id.*, 131). Esto implica para Nagel, aunque no lo ponga en estos términos, que las mentes que el darwinismo puede explicar no tienen verdadera intencionalidad y no pueden ser el asiento de la razón que la filosofía y el lenguaje común analizan con el concepto de verdad. "Verdad" deja de ser un concepto fundamental para analizar lo que es una mente, y es sustituido por "utilidad adaptativa". Contra el darwinismo, Nagel propone entonces un enriquecimiento de los elementos fundamentales del orden natural. Aunque su propuesta es algo ambigua, en su forma más clara sugiere añadir leyes fundamentales que expliquen que la evolución de la mente no es un mero accidente; leyes que de alguna manera impliquen que la capacidad del universo para "generar organismos con mentes" (*Id.*, 132) es un rasgo fundamental del mismo: "[...] la posibilidad de la racionalidad podría ser un rasgo fundamental del orden natural." (*Id.*, 138).

Pero con esta propuesta, la posición de Nagel sobre las relaciones entre explicación y justificación se vuelve insostenible. ¿No es lógico pensar que su explicación más completa logra también justificar y legitimar su concepto de razón? Esta consecuencia es inevitable. Pues si existen leyes que nos pueden explicar cómo ha llegado a existir la razón tal y como la concibe el realista de la razón, entonces no tenemos por qué dudar ya de que la razón tiene las cualidades que el realista le atribuye. La explicación que Nagel propone ofrece también una justificación. Tampoco puede mantenerse el veredicto de imposibilidad que Nagel hace pesar, al concluir su libro, sobre el proyecto naturalista de proporcionar «una comprensión del mundo que se cierre sobre sí misma al incluirnos a nosotros y a nuestros métodos de pensamiento dentro de su alcance» (*Id.*, 143). Habría aquí circularidad, pero la circularidad es inevitable y se la enfrenta mejor desplegando la coherencia de todas las presuntas evidencias racionales con una explicación del mundo lo más comprehensiva posible, que simplemente aceptando como definitivo lo que se nos presenta como evidente.

Con esto queda al menos eliminada la presunta irrelevancia de la explicación para la justificación, la última de las cuatro posibilidades mencionadas al comienzo. Nos quedan entonces las tres primeras, que en el fondo son dos: que la explicación biológica socave y mine los cimientos de la moral tradicionalmente entendida, o que los reafirme y confirme. Esta última posibilidad es, a su vez, subdividida en dos: que

la explicación biológica afiance las más exigentes concepciones de la moral, o sólo las más modestas.

3. La biología como socavamiento de la moral

Siguiendo lo que parece ser su versión oficial de lo que una explicación puede o no puede ofrecer, pero que en realidad solo atañe a la explicación darwinista que él ve ya defectuosa como explicación, Nagel va a concluir, después de un breve argumento, que la explicación darwinista podría a lo sumo socavar los cimientos de la moral, de ningún modo justificarla. Su argumento se puede reproducir de este modo. Supongamos que la investigación biológica demostrase que la disposición humana, nada inusual, a discriminar por razones raciales, es una disposición innata que tenemos por diseño evolutivo. Dada, por ejemplo, una escasez crítica de recursos, negamos –según esta explicación, por una disposición innata– a individuos de otras razas el tratamiento moral y los derechos que concedemos a los de nuestra raza. La disposición evolutiva innata consiste en que la justificación “es de otra raza” para la conducta discriminatoria se nos presenta como definitiva y evidente. Nagel pregunta si eso exime a esa disposición de una crítica moral. Si aceptamos que se la puede someter a crítica moral, como es obvio que lo aceptamos, entonces implícitamente aceptamos que las explicaciones biológicas de nuestras disposiciones morales no son justificaciones.

Qué pasa, continúa Nagel, si encontramos una explicación evolutiva para los pensamientos (por ejemplo, el de la imparcialidad como criterio moral) con los que criticamos la disposición a discriminar moralmente por razones raciales? En este caso Nagel sólo ve dos posibilidades, que son como dos caras de la misma moneda: la crítica del racismo pierde para nosotros su validez racional, o la crítica se mantiene, pero no por su peso racional, sino simplemente porque los pensamientos sobre la imparcialidad son más fuertes que los que nos inclinan al racismo, aunque en sí “no tienen más validez objetiva que el gusto por el azúcar” (*Id.*, 141). La implicación de esta reflexión es clara: si la validez objetiva que creemos ver en el criterio de la imparcialidad se reduce a que estamos biológicamente diseñados para actuar según este criterio, entonces, según Nagel, la validez objetiva es una ilusión. Al comparar nuestro acatamiento de la imparcialidad con el gusto por el azúcar, Nagel juega con la idea de que los gustos son subjetivos. La subjetividad de los gustos significa que, si quisiéramos expresar el agrado que encontramos en el azúcar diciendo “el azúcar es bueno”, inmediatamente tendríamos que reconocer que ese juicio no atribuye ninguna bondad objetiva al azúcar, sino que sólo expresa el efecto subjetivo que produce en nosotros. Del mismo modo, si la disposición a aceptar el juicio “la conducta imparcial es moral-

mente buena” se explica biológicamente, ese juicio pierde su peso racional y objetivo. Pero aquí es donde se observa que Nagel no va al fondo de las cosas. Precisamente en un caso como el gusto por el azúcar, la biología evolucionista puede contar una historia que explica su existencia. Esta historia le quita al término “subjetivo” aplicado a este caso su connotación de arbitrariedad. El placer subjetivo, expresado en el juicio “el azúcar es bueno”, está en función de la necesidad objetiva que nuestro organismo tiene de ella. En otras palabras, el estado mental expresado en ese juicio es una adaptación y cumple una función biológica.

Parece obvio que hay una diferencia entre aceptar un juicio evaluativo porque es objetivamente correcto y aceptarlo porque estoy biológicamente programado para aceptarlo. Sin embargo, no parece implausible interpretar el juicio implícito en nuestro gusto por el azúcar: “el azúcar es bueno”, como un juicio que atribuye valor objetivo: si estamos biológicamente programados para que nos guste el azúcar, lo estamos por buenas razones, y son esas buenas razones las que reconocemos implícitamente al sentir un gusto inmediato por el azúcar.

Esta breve reflexión tiene que ser aún cuidadosamente deletreada para que pueda neutralizar el prejuicio filosófico del que se alimenta Nagel. Pero su propósito no es zanjar la cuestión, sino llamar la atención sobre la lógica de la explicación biológica acerca de por qué un pensamiento parece evidente. El concepto de función y el de adaptación desempeñan allí un papel fundamental, respecto de la que cabe preguntar si no contiene elementos apropiados para una justificación. Nagel en ningún momento menciona esta lógica como un asunto que merezca la reflexión del filósofo. Parte de esta necesaria, aunque ausente, reflexión filosófica sobre el concepto de función y su relación con los conceptos de necesidad e interés es lo que quiero intentar antes de pasar a exponer propuestas de explicación biológica con capacidad justificativa. Se trata de una reflexión que debe tenerse como un telón de fondo sobre el cual esas propuestas deben entenderse.

4. La racionalidad implícita en las funciones biológicas

Simplificando, lo que habría que mostrar es que el concepto de fin o función biológica expresa una forma de racionalidad. Ello abriría una puerta para pensar en la posibilidad de hablar de justificación en este contexto. Empecemos notando que hay que evitar la tentación de decir que la teoría de la selección natural elimina los fines y las funciones.¹ Si entendemos estos conceptos como esencialmente ligados a las intenciones de una mente creadora, entonces es correcto decir que fueron eliminados por la teoría darwinista. Pero Darwin no eliminó la

¹ Esta interpretación se presupone en muchos casos. Cf. Botero 2000: 77.

finalidad, sino que le dio una explicación diferente a la dada por la concepción tradicional basada en la mente creadora. Es cierto que al explicar algo de otra manera lo vemos también con otros ojos, pero eso no hace que no sea lo mismo que veíamos antes, de lo contrario habría que decir que toda nueva explicación, por ser nueva, ya no explica lo mismo. El punto de Nagel es, sin embargo, que Darwin eliminó la finalidad como rasgo *fundamental* de la naturaleza. Me parece que aquí Nagel interpreta erróneamente la teoría de la evolución. Según la explicación de Darwin, *una vez dadas* las entidades que se autoreplican, la finalidad no es accidental, es necesaria, como veremos en un momento. El carácter accidental de la finalidad se tendría que apoyar en algo que está allende la teoría de la evolución por selección natural. Si, como usualmente se entiende, el complemento ideal de la teoría de la evolución es una explicación del surgimiento de moléculas autoreplicantes como un puro accidente a partir de moléculas inertes, entonces la finalidad también existiría accidentalmente en el universo. Pero en realidad, no hay nada en la explicación darwinista de la finalidad que implique necesariamente un surgimiento puramente accidental de las moléculas autoreplicantes. Una vez dadas las moléculas que se autoreplican, la finalidad se deriva según la explicación darwinista como algo inevitable por la selección natural.

El concepto de función y de fin se puede aplicar a los rasgos o propiedades de moléculas autoreplicantes de la siguiente manera. La función o fin de un rasgo es el efecto de ese rasgo, cuando dicho efecto tiene consecuencias beneficiosas para la replicación de la molécula y explica por qué ella sobrevive y se replica con más éxito que otras que no exhiben ese rasgo. Es importante que dicho rasgo, aun habiendo llegado a la existencia por un proceso causal determinado, no lo ha hecho *con el fin* de tener consecuencias positivas sobre la supervivencia y replicación de la molécula. Por otro lado, ese rasgo y sus efectos tienen que marcar realmente una diferencia en el proceso causal que hace que las moléculas que lo poseen se repliquen más que las que no lo poseen. Si todas las moléculas se replicaran por igual sin importar cuáles fueran sus rasgos –debido, por ejemplo, a una inverosímil abundancia de recursos en el entorno– no tendría sentido hablar de funciones, ni de adaptaciones, ni tampoco de diseño biológico. La reproducción diferencial heredable en una población de moléculas autoreplicantes es una condición suficiente para que sus rasgos tengan funciones. Una vez dada esta condición, las funciones se dan como algo necesario en ese contexto. La reproducción diferencial heredable es aquello en lo que termina toda explicación funcional. Puede por ello verse como la *función o fin último*, si se tiene cuidado en evitar las connotaciones tradicionales de este concepto que son totalmente impropias en el contexto darwinista.

Precisamente debido a estas connotaciones se ha argüido que, al

considerar la reproducción como el fin o función final de toda entidad biológica, no estamos expresando ningún hecho natural, sino sólo situando los hechos en contextos subjetivos, o relativos a un observador (cf. Botero 2000: 76-7). Si yo valoro un estado de cosas E, E es para mí un fin y un propósito. Por mi parte, no estoy de acuerdo con que el darwinista esté introduciendo, de manera implícita, a la reproducción o supervivencia como un valor a promover. Si optamos por decir que la reproducción diferencial heredable en una población de entidades biológicas es la función o fin último, sólo estamos diciendo que, sin necesidad de que una mente valore la reproducción y la ponga como un fin, la reproducción diferencial hace posible que algunos rasgos de esas entidades existan por los efectos que producen. La reproducción diferencial heredable alude a un conjunto de circunstancias objetivas que, no importa si son valoradas o no por alguna mente, conducen a que sea un hecho real y natural que las entidades que se replican en esas condiciones tengan rasgos que existen por los efectos que producen. Que un rasgo exista por causa de los efectos que produce es el elemento clave en la noción de función, pues se trata de un aparente contrasentido desde el punto de vista de la noción de causa que se maneja en física, que explica lo que existe por las causas antecedentes y no por los efectos. Dada una mente diseñadora, es fácil explicar el aparente contrasentido, pues es la *representación* de los efectos (en la mente) lo que realmente causa los rasgos que, aparentemente, se explican por sus efectos. El mérito de Darwin es precisamente haber aclarado, sin recurrir a la idea de la mente diseñadora, un proceso natural que también elimina el aparente contrasentido.² Después de Darwin, ya no tiene fuerza objetar que, dado que la reproducción diferencial no es realmente un fin de nadie, ella no puede fundar y legitimar ninguna atribución de fines o funciones. El argumento es una *petitio principii*, pues supone que tengo que explicar funciones o fines de objetos que no son mentes, apelando a los fines (intenciones) de una mente.

Defendido ya el concepto de fin y función en el marco de la teoría de la evolución, no es difícil legitimar el uso de los conceptos de necesidad e interés objetivo. En el contexto de la biología, el concepto de necesidad, en su sentido más amplio, se puede aplicar a todo aquello que tiene una relación causal eficiente con la supervivencia y reproducción de un organismo. Un organismo tiene necesidad de X, cuando X promueve o conduce causalmente a su supervivencia y reproducción. (Para mayor generalidad, X puede entenderse como un estado de cosas.) En un sentido más estricto, las necesidades de un organismo se vinculan directamente con sus funciones, de modo que el

² Un artículo clásico que analiza de este modo el concepto de "función" es: Wright 1973.

organismo tendrá necesidad de un subconjunto de los estados de cosas que conducen a su supervivencia y reproducción, a saber el subconjunto que es expresamente objeto de sus funciones. En este sentido puede decirse que la primera necesidad de un organismo es el ejercicio libre y sin obstáculos de sus funciones. El concepto de interés objetivo es intercambiable con el de necesidad. Si con base en estas reflexiones sobre el concepto de función biológica se concede que los intereses de cada organismo existen desde un punto de vista biológico y que están reflejados en sus funciones, se concede también que el "espacio lógico de las razones" (prácticas) tiene su fundamento en la realidad biológica.

5. La sociobiología y las tradiciones rivales en la psicología moral

Los organismos biológicos tienen necesidades e intereses objetivos y tienen funciones destinadas a satisfacerlos. Las necesidades e intereses objetivos están siempre "reflejadas" en las funciones y actividades del organismo. Pero esto no implica que las funciones "representen", en el sentido intencional, las necesidades. Bacterias marinas para las cuales el oxígeno es letal, tienen la necesidad de evitarlo. Para ello han desarrollado un dispositivo complejo con la función de detectar la dirección en la que se encuentran altas concentraciones de oxígeno y de evitar que la bacteria nade en esa dirección. Esta función refleja su necesidad, pero suponemos con buenas razones que estos dispositivos en la bacteria no implican una representación del oxígeno como algo que hay que evitar, es decir, no pueden equipararse a un deseo de evitarlo. Los deseos, en tanto actitudes proposicionales, requieren capacidad conceptual, ausente en la bacteria. En su caso hablaríamos más bien de un instinto o de un mecanismo. Aun así, la representación subjetiva que asciende en diversos grados desde el mero instinto o "reflejo funcional de un interés" hasta la plena autoconciencia o reconocimiento cognitivo del mismo, no puede ser sino un producto de la selección natural, ligada a la evolución del sistema nervioso y del cerebro. Los deseos pueden valer como una primera representación de la necesidad de un organismo y constituirían el comienzo de la razón práctica. Aunque aun no sabemos si es legítimo atribuir deseos a los animales, su existencia en los humanos nos permite sostener plausiblemente que se trata de una capacidad que tiene un valor adaptativo, siempre y cuando represente las necesidades del organismo con fidelidad y confiabilidad.

Para un filósofo, parece ser más fácil aceptar la relevancia de la teoría de la evolución para los animales, y en general para la naturale-

za apetitiva o animal, que aceptarla para las capacidades conceptuales. En esta actitud operan intuiciones poco confiables sobre la naturaleza supra-natural de la razón. Las mismas se movilizan al evaluar filosóficamente el impacto de las nuevas ideas de la biología del comportamiento en la filosofía moral. Esto es inevitable, porque las teorías evolutivas del comportamiento social han dicho poco sobre los mecanismos psicológicos que lo sustentan.³ Los filósofos deben pues completar este silencio con sus propias intuiciones o teorías sobre la estructura motivacional que sustenta el comportamiento moral y parecen haber encontrado un consenso espontáneo, aunque por razones diversas, en referir los comportamientos explicados por la sociobiología a mecanismos de índole emotiva y apetitiva. Los que en la tradición de Hume basan la moral en emociones y sentimientos tienen menos resistencia a admitir la relevancia moral de las teorías sociobiológicas. Ese es el caso de John Mackie, cuya adhesión a esta tradición es conocida desde su libro *Ethics. Inventing Right and Wrong*. Quienes, en cambio, se sitúan en la tradición de Kant y piensan que la razón es el asiento de la capacidad moral, sostienen que las teorías sociobiológicas –asumiendo también que son sólo válidas respecto de la naturaleza emotiva– no son suficientes para explicar y fundamentar la moralidad. Como ellos quieren, sin embargo, permanecer dentro de la perspectiva evolutiva, piensan en términos de un modelo más complejo. Sostienen que la moralidad no es una adaptación directa, explicable sociobiológicamente, sino un producto derivado y evolutivamente inesperado. Es el caso del modelo esbozado por Colin McGinn, y continuado por Peter Singer.

6. McGinn y la imposibilidad evolutiva del altruismo incondicionado

El objeto de la explicación sociobiológica ha sido principalmente el comportamiento; muy poco ha dicho sobre los mecanismos psicológicos que lo sustentan. Adoptando explícitamente la perspectiva del gen como unidad de selección y haciendo uso de ideas afines a la teoría de juegos, la sociobiología ha explicado dos tipos de comportamientos altruistas, que se pueden denominar altruismo nepotista y altruismo recíproco. Suponiendo que un gen o grupo de genes codifique comportamientos altruistas –que ceden beneficios a individuos distintos del agente altruista y sus descendientes– es predecible que, si la donación de beneficios es indiscriminada, esos genes desapare-

³ Trivers (1972) presenta unas especulaciones esquemáticas sobre la función de emociones como la gratitud y la culpa, pero no alcanzan para explicar su verdadera contribución a la estrategia racional de la reciprocidad. Al respecto véanse las críticas de Frank 1988, cap. 2.

cerían tarde o temprano del *pool* genético. Si en cambio el altruismo se dirige a parientes cercanos, el gen que codifica el altruismo se beneficiaría también a sí mismo, dado que hay una probabilidad alta de que estos parientes también tengan los genes altruistas. Lo mismo cabría decir si el comportamiento altruista se dirige a otros individuos altruistas que de algún modo se reconocen como tales (por ejemplo, por su comportamiento). La ley de la selección natural operando al nivel de los genes pone, pues, condiciones al comportamiento altruista que se puede explicar como adaptación biológica. Eso es lo que Dawkins dio en llamar el “egoísmo” de los genes: ellos “buscan” beneficiarse a sí mismos, es decir, su propia replicación. Y aunque es natural suponer que las condiciones bajo las cuales el comportamiento altruista puede ser adaptativo se reflejan también en los mecanismos psicológicos subyacentes, la teoría del “gen egoísta” no puede *a priori* excluir motivaciones altruistas, pues las condiciones valen para el comportamiento a sus consecuencias objetivas, y nada impide en principio que motivaciones altruistas promuevan más eficientemente el “egoísmo” de los genes.

La pregunta por el espacio dejado por la sociobiología para el altruismo como motivación está en el centro de ambos modelos mencionados arriba.⁴ Pues para ambos la naturaleza de la moralidad gira en torno al altruismo motivacional, sus alcances y sus límites. Para McGinn la moralidad y los genuinos ideales morales son “radicalmente no-egoístas”, “altruismo puro y desinteresado”. Esta concepción de la moralidad en McGinn es lo que decide su juicio negativo respecto de la posibilidad de explicar sociobiológicamente la moralidad. Pues lo que la sociobiología ha dejado muy en claro, es que no puede haber selección de genes que determinen un comportamiento altruista en este sentido, un altruismo puro, desinteresado y radicalmente no-egoísta, que no ponga condiciones de ningún tipo (por ejemplo, la condición de que el beneficiario sea un miembro de la familia, o la condición de recibir algo equivalente a cambio). Un altruismo “incondicionado” no podría nunca ser adaptativo. Hay que destacar que la posición de McGinn no tiene nada que ver con un rechazo de una visión biológica y genéticamente determinista del ser humano. El punto es que la sociobiología ha demostrado que si existiese un gen para el altruismo sin condiciones, que se dirigiese por igual a todos los individuos de la misma especie –o incluso a todo organismo con intereses–, sería rápidamente eliminado por la selección natural.

Hay otra línea de argumentación que no está explícita, pero que opera en el transfondo, y que se apoya en las ideas de Nagel sobre el altruismo en su *The Possibility of Altruism*. Nagel defendió en esta obra que el altruismo no es originalmente un sentimiento, sino una capaci-

⁴ La exposición que sigue se basa en Mackie 1982 y McGinn 1979.

dad racional de reconocer los intereses de otras personas como razones también válidas para mí, es decir, de ser motivado directamente por el reconocimiento de esos intereses, sin que tenga que haber un deseo benevolente de por medio. McGinn piensa que el mecanismo que está detrás de la moralidad es cognitivo en este sentido nageliano. Pero asume además que hay un altruismo emotivo-apetitivo, y sostiene que es éste el que la sociobiología tiene por objeto. Este altruismo es algo de lo que los animales también son capaces, y que “no tiene la etiología psicológica requerida para ser catalogado de moral” (McGinn 1979: 89). (McGinn atribuye deseos a los animales al tiempo que les niega, a mi parecer inconsistentemente, capacidades conceptuales.) En otras palabras, el altruismo propiamente moral se funda en un mecanismo cognitivo, que reconoce la validez objetiva de los intereses de los demás sin poner condiciones, como la de ser familiares, amigos, socios, o miembros de la misma comunidad política, legal o étnica; y la sociobiología no puede explicarlo, por estar limitada a mecanismos apetitivo-emocionales.

McGinn acepta acriticamente el viejo prejuicio de que la biología tiene jurisdicción sobre la naturaleza animal de los deseos y apetitos, pero no sobre la razón. En su caso, empero, la presencia de este prejuicio sorprende. Pues él no es de los filósofos que piense que la dignidad de la razón precluye un origen genético. Él mismo dice que si hubiese un gen para el altruismo cognitivo incondicionado, para ser motivados por los intereses de otras creaturas con sólo reconocerlos cognitivamente, entonces no tendría reparos en aceptar que la moralidad es una dotación genética. (*Id.*, 90) No es la dignidad de la razón lo que impide que esto sea así, sino las razones aducidas por la sociobiología, según la cual un gen para el altruismo incondicionado no es evolutivamente viable. Pero la tesis de la moral como adaptación soportada genéticamente se torna plausible, si abandonamos la tesis que equipara motivación moral con altruismo incondicionado. Nada nos impide en realidad concebir las capacidades motivacionales de la razón de modo que apoyen una moral más modesta y un altruismo condicionado al estilo sociobiológico. Pero McGinn es ciego a esta posibilidad, y en ello, creo yo, tiene su parte el prejuicio que lo obliga a juzgar que, si la sociobiología puede explicar una moralidad basada en un altruismo condicionado, entonces ese altruismo no puede basarse en un mecanismo cognitivo-racional, ni tampoco puede ser el verdadero núcleo de la moral.

7. La moralidad como producto derivado

Examinemos ahora cómo explica McGinn “la etiología psicológica” del altruismo incondicionado. La idea básica es que percibir objetivamente los intereses de otras personas (o creaturas en general), equiva-

le a asumirlos como razones válidas, como motivaciones en nuestras deliberaciones prácticas. La capacidad para percibir el mundo objetivamente, incluidas las otras mentes, fue seleccionada por sus inmensas ventajas. Sucede, empero, que esta capacidad implica, quíerese o no, la capacidad de percibir los intereses de otras creaturas como razones objetivas, en el sentido de ser válidas allende los límites del individuo. La capacidad para el altruismo incondicionado es una consecuencia inevitable e inesperada, un producto derivado de la capacidad de conocer el mundo. Esta explicación permite responder a la pregunta acerca de por qué el altruismo sin condiciones no ha sido proscrito por la selección natural. La diferencia entre las enormes ganancias de la representación objetiva del mundo y los costos comparativamente menores de la moralidad, y el hecho de que ambas sean indesligables, es lo que explica que la selección natural no haya podido eliminar la moralidad.

Esta explicación de la evolución de la moralidad pese a sus costos depende, como es obvio, de que la capacidad de ver el mundo objetivamente implique en realidad la capacidad de reconocer los intereses de otras creaturas como razones válidas, en sentido práctico, para todo agente deliberante. El mismo Nagel no logró argumentarlo convincentemente. Sin duda que la capacidad de percibir el mundo objetivamente me lleva a reconocer a otras creaturas como seres racionales o sentientes, con intereses que son para ellos tan válidos como los míos lo son para mí. No es de ningún modo obvio, en cambio, que este sentido teórico de objetividad implique una objetividad práctica, es decir, una capacidad de motivar a agentes distintos al sujeto original del interés. McGinn pasa por alto la diferencia entre reconocer intereses como rasgos objetivos de la psicología de una persona, y reconocer la validez objetiva práctica de intereses ajenos en las deliberaciones de la persona que los reconoce como ajenos. Hay circunstancias en las cuales conocer objetivamente los intereses ajenos nos sirve para prever el comportamiento de aquellos a quienes queremos vencer estratégicamente. Sin explicar cómo se pasa del conocimiento teórico al reconocimiento práctico, la explicación de la evolución del altruismo incondicional en McGinn no logra convencer.

Por otro lado, la sola tesis nageliana de la objetividad de las razones es más sutil y compleja de lo que McGinn sugiere y eso sin ver aún el argumento en su favor. Con buenas razones podría dudarse de si la tesis de la objetividad en Nagel es apropiada para defender un altruismo incondicionado. Pero, en todo caso, la sola capacidad de ver el mundo objetivamente no basta para que los intereses de otra persona puedan ser vistos por mí como razones objetivas en el sentido de reclamar validez en mis propias deliberaciones. Eso significa que McGinn no logra ofrecer una historia convincente de la "etiología psicológica" del altruismo que él llama puro, desinteresado, radicalmente no-egoísta (a lo cual yo agregué el apelativo de incondicionado).

Puede que el proyecto de apoyar la moralidad en un altruismo de esta naturaleza no sea una buena idea. Si el de altruismo moralmente relevante contiene necesariamente elementos de condicionalidad, esto le abre de nuevo las puertas a la sociobiología y a su concepto de altruismo. La atracción que el modelo de McGinn ejerce sobre la filosofía se debe, en mi opinión, a que hace un espacio para la vieja idea de la autonomía de la razón práctica frente a lo meramente animal y controlado por los genes, en un contexto que se confiesa, no obstante, partidario de una visión evolucionista del ser humano. Otro ejemplo claro de esa atracción es el excelente libro de Peter Singer sobre la sociobiología y la ética: *The Expanding Circle. Ethics and Sociobiology* (1983). La tesis fundamental que Singer desarrolla es la tesis ya esbozada por McGinn: que si bien la razón es una adaptación con base genética, no lo es en relación con sus capacidades morales. Estas son una consecuencia inesperada de la dinámica interna y autónoma de la razón. Singer describe el surgimiento de la moral, aun en sus formas irreflexivas y pre-modernas –el sistema de costumbres gobernando una tribu, por ejemplo– como un producto de esta razón autónoma: “la transformación del comportamiento social programado genéticamente en costumbres sociales significó la primera limitada aplicación de la razón a lo que hasta entonces había estado bajo el dominio indisputado de los genes” (*Id.*, 94-5). La implicación cuestionable de la cita, es que el comportamiento social programado genéticamente, el altruismo hacia los parientes cercanos y el altruismo recíproco, no supone capacidades racionales. Esto no es consistente con la forma como Singer entiende el concepto sociobiológico del altruismo recíproco (*cf. Id.*, 37-49), a saber, como implicando capacidades cognitivas. Pasajes como ese testimonian la atracción ejercida por el viejo prejuicio de que la razón (al menos la razón práctica) está por encima de lo animal y de lo que los genes pueden construir. Opino, en cambio, que el verdadero reto para la recepción filosófica consiste en recuperar para la filosofía moral el altruismo condicionado por la reciprocidad propuesto por la sociobiología, sin precluir la posibilidad de que se trate de una motivación apoyada en capacidades cognitivas o racionales.

8. Mackie: La reciprocidad y los límites del altruismo

El modelo de recepción filosófica de la sociobiología elaborado por Mackie puede proporcionar, sin quererlo, algunas ideas para este proyecto. Digo sin quererlo, porque Mackie se reclama partidario de la tradición moral proveniente de Hume, precisamente la tradición a la que se opone la idea nageliana del altruismo como capacidad cognitivo-racional. En la línea de Hume, la moralidad, y lo que de altruismo

haya en ella, se concibe como una capacidad fundada en sentimientos y emociones. También para Mackie la sociobiología queda limitada a explicar comportamientos apoyados en mecanismos motivacionales exclusivamente emotivos, sin que ello le quite relevancia moral. Mackie afirma: "Inicialmente, se explica de esta manera [sociobiológica] el comportamiento retributivo; pero en creaturas con un alto nivel de conciencia habrá naturalmente sentimientos retributivos, tal como la ira y el resentimiento que acompañan y sustentan el comportamiento." (Mackie 1982: 155). Aunque Mackie no concede ningún papel prioritario a mecanismos de tipo cognitivo, eso no implica que los mecanismos emotivos aludidos carezcan de racionalidad. Pues Mackie ve como esencial a la sociobiología su capacidad para explicar por qué una benevolencia (altruismo) universal o incondicionada no es racionalmente recomendable. No se trata solamente de que una moral de este tipo tendría que luchar contra tendencias emotivas genéticamente determinadas, sin importar si ellas son racionales o no. Mackie hace lo que Nagel recomienda debe hacerse frente a las disposiciones morales explicadas biológicamente: someterlas a la crítica racional (ver sección 3. de este ensayo). Se pregunta si los mecanismos emotivos se encuentran dentro de lo que es racionalmente recomendable. Las disposiciones biológicas que Nagel quería entender como hechos brutos sin racionalidad, tienen una lógica que las repalda y las explica como adaptaciones.

Se trata de que la moral de la "benevolencia universal" es vulnerable. Mackie se opone a la moral utilitarista y a la kantiana, señalando que estas propuestas morales, en la medida en que están "dirigidas incondicionalmente al bienestar de toda una comunidad" (*Id.*, 162), o más bien de la humanidad en cuanto tal, implican la prescripción incondicional de tratar a los demás como nos gustaría que nos traten, sin tener en cuenta si los demás harán lo mismo con nosotros. Independientemente de si Mackie interpreta aquí correctamente al kantismo y al utilitarismo, esta comprensión de la moral la hace vulnerable a la explotación. Mackie está dando aquí una razón en favor del altruismo recíproco, que se puede ilustrar con ayuda de la teoría de juegos. En el mundo animal y humano el conflicto de intereses entre individuos nunca es total y los juegos de estrategia más importantes para modelar las presiones selectivas son los de conflicto parcial, en los que hay también una confluencia parcial de intereses. El ejemplo clásico es el 'dilema del prisionero'. La mejor estrategia en este juego tiene un componente cooperativo. Pero una cooperación *incondicional* no arroja ganancias: está abierta a la explotación. Los genes que la determinan, son explotables, invadibles y terminan siendo desplazados por genes que prosperan a costa suya. La estrategia más exitosa tendrá un elemento adicional: la capacidad de ejercer represalias contra cualquier forma de agresión y explotación. Una estrategia exitosa debe protegerse contra las pérdidas de la explotación, y aprovecharse

del beneficio, vedado a los egoístas, de los frutos de la cooperación.⁵ Los genes que la controlan serán también los que más se repliquen.

Esta lógica de las utilidades es la lógica que realmente tiene el proceso selectivo, donde las utilidades están representadas por la replicación diferencial que arrojan las diferentes estrategias. Obviamente, los genes no conocen esta lógica, ni tampoco la mayor parte de los animales que los genes construyen. Mackie afirma que los contendores animales que se enfrentan en estos juegos no tienen que elegir atendiendo a esta lógica: "no hay necesidad de cálculo o elección deliberada...; el mecanismo de la selección natural mimetiza la finalidad produciendo comportamientos instintivos que se asemejan a los que resultarían de un cálculo inteligente" (*Id.*, 154). En esta misma línea de razonamiento señala que en el caso de los humanos o de animales con alto grado de conciencia, habrá también sentimientos retributivos como la gratitud y el resentimiento. Esos sentimientos son seleccionados en el curso de la evolución por sus ventajas adaptativas, mediante una selección de los genes que los determinan, y sin que los organismos tengan que reconocer explícitamente las razones por las cuales esos genes y los sentimientos que ellos programan fueron seleccionados. Mackie recalca que los mecanismos seleccionados –ya sea social o evolutivamente– para sustentar la conducta moral pueden ser inclinaciones espontáneas, y no necesariamente principios o propósitos conscientemente calculados (*Id.*, 158).

Esto se puede entender de dos maneras, y creo que Mackie tiene ambas en mente, pero sólo una me parece correcta. La correcta es que la disposición a actuar cooperativamente no tiene que deberse a un cálculo o a un mecanismo de tipo egoísta. El mecanismo seleccionado para la cooperación bajo reciprocidad lo fue por su capacidad de producir un comportamiento que promueve los intereses del individuo y en última instancia de los genes que lo construyen. Pero eso no significa que el mecanismo tenga esos intereses a la vista como su fin último (*Id.*, 161). La terminología de Mackie, empero, también sugiere que la razón no tiene ningún papel, ni siquiera en los humanos, en el mecanismo que sustenta de las actitudes cooperativas. Pues él usa los términos "espontáneo" y "no-calculado" (instintivo) para lo que está genéticamente determinado sin intervención de la inteligencia como mecanismo próximo (*cf. Id.*, 154, 156).

El punto ahora en discusión es si una lógica racional, no necesariamente egoísta, es separable de los mecanismos que actualmente sustentan el comportamiento cooperativo y altruista en los humanos. Es posible que haya organismos capaces de ejercer represalias contra la agresión gracias a mecanismos puramente instintivos sin capacidades

⁵ La más célebre de estas estrategias es la denominada TIT FOR TAT. *Cf.* Axelrod & Hamilton 1981.

conceptuales. Más difícil me parece ya sostener que tales conductas, si se apoyan en sentimientos como el resentimiento, no vayan necesariamente acompañadas de y complementadas por esas capacidades, pues no puedo resentirme contra acciones y personas, si no es bajo una determinada descripción de las mismas. Y me parece plausible que esta descripción debe contener al menos un asomo de conciencia de la lógica que explica la existencia y la función de esas emociones, más aun si ellas se insertan en un contexto de interacción social y, específicamente, de intercambios cooperativos. La descripción e interpretación de los mecanismos psicológicos subyacentes es importante para la recepción de las ideas sociobiológicas y debe llevarse un poco más allá de lo que lo han hecho los modelos aquí examinados.

Quiero dar una idea de la interpretación psicológica requerida, sirviéndome de una historia que Mackie cuenta acerca de cómo pudo haberse originado la convención del trueque regular de bienes. La historia cuenta cómo pudo surgir esa convención antes de que existieran la institución del contrato y otras instituciones sociales afines (promesas, acuerdos etc.). Se trata del surgimiento de la forma más básica de cooperación, la que pudo haber dado origen a primeras formas de sociedad por fuera de la familia. En este sentido, la historia presupone sólo capacidades y disposiciones psicológicas básicas, las que, por no tener que apoyarse en instituciones sociales preexistentes (salvo la familia), indican que es lo que la naturaleza humana debe ya tener, lo que la evolución nos debe ya haber dado, para poder involucrarnos en empresas cooperativas. Se parte de dos grupos familiares que se especializan en producir bienes diferentes. Ambos podrían satisfacer necesidades propias con el bien que produce el otro y ambos lo saben. Pero no tienen razones para confiar uno en el otro. Aun así, uno de ellos arriesga un acercamiento y ofrece un obsequio. Si el otro lo acepta sin retribuir apropiadamente, el primero ejerce represalia, no repitiendo el obsequio o reduciéndolo. Si esta dinámica se repite, respondiendo obsequios con obsequios y abusos con represalias, el resultado es el surgimiento de un hábito o convención de trueque de bienes. La historia es simple de entender, pero es incompleta mientras no se diga qué capacidades psicológicas se requieren para explicar el comportamiento de los sujetos en esta historia. Obviamente, ambos sujetos o familias deben percibirse como sujetos con intereses y necesidades que pueden ser satisfechas por los bienes que ambos producen. Si nos concentramos en el sujeto que arriesga el acercamiento, las preguntas claves se refieren a cómo explicamos que se decida a arriesgarse y que ejerza represalias si su obsequio no es correspondido. ¿Cómo hemos de entender este comportamiento que asume un riesgo inicial, respaldado, si es necesario, con represalias posteriores?

Para la interpretación psicológica adecuada de este comportamiento me parece importante separar el concepto de cálculo del concepto

del egoísmo. Quien asume el riesgo inicial, tiene que saber que la respuesta del otro puede ser positiva o negativa, y en el segundo caso anticipar la necesidad de adoptar una actitud punitiva (puede ser en ello empujado por el resentimiento, pero no me parece que sea indispensable que así sea, a no ser que se llame resentimiento también a la percepción o comprensión de que en ese caso es adecuado reponder con represalias). También es natural suponer que el riesgo se asume cuando se cree que hay al menos la misma probabilidad de una respuesta positiva como de una negativa. En ese sentido, es adecuado pensar que hay un cálculo de por medio. Quien asume el riesgo ya ha calculado las alternativas del otro jugador y las propias jugadas posteriores. Pero denominar egoísta a este cálculo es restringir innecesariamente los motivos involucrados. Obviamente, el cálculo de las respuestas y contrarrespuestas implica no sólo una conciencia de los propios intereses, sino la disposición a hacerlos valer en la interacción. Pero no hay razón para negar que la misma disposición existe respecto de los intereses de la otra parte. El que arriesga el acercamiento inicial, no sólo tiene una conciencia de ellos o de su existencia; precisamente al arriesgarse, el sujeto pasa de esta conciencia meramente teórica a un reconocimiento práctico de los mismos. Al dejar el obsequio, el sujeto está reconociendo la objetividad, en sentido nageiliano, de los intereses de la otra parte. Por otro lado, en el caso de que la otra parte no retribuya, la represalia no es otra cosa el reconocimiento de la objetividad de los *propios* intereses. No ejercer represalias es no respetarse a sí mismo y a los propios intereses, y negarles la objetividad práctica que se le reconoce a los intereses ajenos. La represalia no es sino el rechazo a la explotación que amenaza la existencia del altruismo. No implica instrumentalización del otro para los propios fines. Al contrario, habría que decir que la exigencia de reciprocidad es el rechazo a la instrumentalización de uno mismo para los fines del otro. Es una condición razonable y no contradice la autenticidad del altruismo.

La recepción filosófica de la sociobiología adolece, pues, de una limitación importante, que consiste en una interpretación prejuiciada de los mecanismos psicológicos que subyacen a las conductas morales de las que habla la sociobiología. El consenso tácito, en los autores aquí examinados, en referir esos resultados conductuales a mecanismos apetitivos, es infundado. Si la reflexión aquí realizada es correcta, la sociobiología puede explicar un altruismo condicionado basado en capacidades cognitivas y, al mismo tiempo, ofrecer elementos para una justificación de esa concepción de la moral. He dejado aquí de lado otras propuestas que pueden interpretarse más bien como propuestas de socavamiento que de justificación de la moral (principalmente Alexander 1987). El propósito de estas reflexiones, presentadas de la mano de la recepción filosófica en un puñado de autores escogi-

dos, fue sugerir que la evolución puede habernos dado la capacidad para un altruismo genuino, para reconocer la objetividad práctica –en el sentido nageliano– de los intereses de los demás. Dicho de otro modo, tenemos por diseño biológico la capacidad de vernos a nosotros mismos y a los demás como fines, y de otorgarnos mutuamente esa calidad en nuestras interacciones y ello nos permite hablar tanto de la explicación como de la justificación biológica de la moral. La diferencia entre esta tesis y la tesis metafísica según la cual somos realmente fines en sí mismos puede ser menos relevante de lo que estamos inclinados a creer. En nuestra naturaleza biológica está la capacidad de otorgarnos mutuamente la categoría de fin y eso basta para darle legitimidad a nuestras actitudes morales.

Bibliografía

- Alexander, R. (1987). *The Biology of Moral Systems*. New York: Aldine de Gruyter.
- Axelrod, R. & William, H. (1981). "The Evolution of Cooperation". En: *Science*, 211: 1390-6.
- Botero, J. J. (2000). "El problema de la representación". En: Botero, J.J., Ramos, J. & Rosas, A. (eds.), *Mentes Reales*. Bogotá: Universidad Nacional de Colombia / Siglo del Hombre: 57-80.
- Dawkins, R. (1976) *The Selfish Gene*. Oxford: Oxford U.P.
- Frank, R. (1988). *Passions Within Reason. The Strategic Role of the Emotions*. New York: W. Norton & Co.
- Mackie, J. (1976). *Ethics. Inventing Right and Wrong*.
(1982) "Co-operation, Competition and Moral Philosophy". En: Mackie, J. *Persons and Values*, Vol.II. New York: Oxford U.P.: 152-9.
- McGinn, C. (1979). "Evolution, Animals and the Basis of Morality". En: *Inquiry* 22: 81-99.
- Nagel, Th. (1997). *The Last Word*. New York: Oxford U.P.
(1970). *The Possibility of Altruism*. Princeton: Princeton U.P.
- Singer, P. (1983). *The Expanding Circle. Ethics and Sociobiology*. Oxford: Oxford U.P.
- Trivers, R. (1971). "The Evolution of Reciprocal Altruism". Reimpreso en: Clutton-Brock, T.H & Paul H. (eds), *Readings in Sociobiology*. San Francisco, CA: Freeman, 1978.
- Voland, E. (1999). "The Nature of Solidarity". En: Bayertz, K. (ed.), *Solidarity*. Dordrecht-Boston-London: Kluwer A.P.: 157-72.
- Wilson, E.O. (1975). *Sociobiology: The New Synthesis*, Harvard: Harvard U.P.
- Wright, L. (1973). "Functions". En: *The Philosophical Review* 82: 139-68.