Revista Colombiana de Filosofía de la Ciencia
Universidad El Bosque
filciencia@unbosque.edu.co
ISSN (Versión impresa): 0124-4620
COLOMBIA

2003
Maximiliano Martínez
LA EVOLUCIÓN DEL ALTRUISMO
Revista Colombiana de Filosofía de la Ciencia, año/vol. 4, número 8-9
Universidad El Bosque
Bogotá, Colombia
pp. 27-42

Red ALyC
LA HEMEROTECA CIENTÍFICA EN LÍNEA EN CIENCIAS SOCIALES
www.redalyc.org

Revista Colombiana de Filosofía de la Ciencia
Vol. 4 • Nos. 8 y 9 • 2003 • Págs. 27-42

La evolución del altruismo

Maximiliano Martínez
Departamento de Filosofía
Universidad Nacional de Colombia
mmartinezb@unal.edu.co

Resumen

Las diferentes tradiciones filosóficas han intentado dar, a lo largo de la historia, una respuesta convincente a la pregunta por el origen de nuestros sentimientos morales. El naturalismo no ha evitado esta empresa. Desde *Descent of Man* de Darwin, la propuesta naturalista* ha ganado tanta evidencia y consistencia como adeptos. Gracias al impulso multidisciplinario que ha tenido en las últimas tres décadas el proyecto naturalista, principalmente desde la biología y la psicología evolutivas, se han podido bosquejar unos lineamientos promisorios en cuanto al origen de la ética, sin necesidad de recurrir a las explicaciones comunes basadas en un origen divino o en propiedades morales *sui generis*. El propósito de este artículo es mancomunar algunas importantes investigaciones (tanto empíricas como teóricas) realizadas desde diversos sectores por pensadores especialistas en el tema. De esta manera se pretende argumentar, desde un marco biológico/evolutivo (apoyado en la teoría de juegos), en favor de la tesis que afirma que nuestros sentimientos morales de cooperación son

* Postura científico-filosófica que sostiene que todos los objetos y eventos del mundo son parte de la naturaleza, incluyendo al hombre en todas sus dimensiones. Por ende, el ser humano es en todos sus aspectos describible por la ciencia.

producto de la evolución por selección natural (en su versión clásica): la función de la moral es permitir la provechosa convivencia pacífica entre individuos por naturaleza egoístas. Se defenderá también que los motivos psicológicos que llevan a realizar una acción moralmente buena deben ser genuinamente altruistas.

Palabras clave

Evolución, moral, selección natural, altruismo, egoísmo, cooperación, dilema del prisionero, tit for tat, Axelrod, Hamilton, Trivers.

Abstract

Throughout history different philosophical traditions have tried to give a convincing answer to the question concerning the origin of our moral feelings. Naturalism has not avoided this matter. Since Darwin's *Descent of Man* the naturalist proposal* has gained as many supporters as it has evidence and consistency. Thanks to the multidisciplinary boost, mainly from evolutionary biology and psychology, which the naturalist project has received over the last ten years, they have been able to outline some promising lines of thought regarding the origin of ethics, without the need to resort to the usual explanations based on divine origin or moral attributes *sui generis*. The purpose of this article is to gather together several important investigations (both empirical and theoretical) from various different sectors, carried out by thinkers specialized in this field. In this way, and from a biological/evolutionary framework (supported by the game theory), this article intends to argue in favour of the theory which affirms that our moral feelings of co-operation are products of the evolution of natural selection (in its classic version): the function of morality is to allow peaceful and advantageous co-existence between individuals who are by

* A philosophical/scientific position which holds that all worldly objects and events are part of nature, including man in all his magnitude. Hence, human beings in all their aspects are describable by science.

nature egoistic. The article will also champion the idea that the psychological motives needed to bring about a good moral action must be genuinely altruistic.

Key words

Evolution, morality, natural selection, altruism, egoism, cooperation, prisoner's dilemma, tit for tat, Axelrod, Hamilton, Trivers.

1. TIT FOR TAT: la supervivencia del altruista

El mismo Darwin había ya notado que el comportamiento altruista constituye un problema para la teoría de la evolución por selección natural.[1] En sentido biológico, podemos definir "altruismo" como el comportamiento que beneficia a otro organismo, no cercanamente relacionado, y que va en aparente detrimento del organismo que se comporta de manera altruista (beneficio y detrimento son definidos en términos de aptitud biológica). Si la selección actúa sobre individuos y no sobre grupos, el comportamiento altruista tendría que conducir a la extinción de los altruistas. ¿Cómo se explica entonces que haya organismos que se comportan de manera altruista?

Si pudiésemos recurrir a la teoría de la selección de grupos (como había sugerido Darwin)[2] el altruismo se explicaría de la siguiente manera. Si bien el comportamiento altruista no beneficia al individuo mismo que lo realiza, sí beneficia al grupo al cual pertenece. Si la competencia entre grupos es mayor que la competencia

---

1 Darwin, en el capítulo 5 del *Origen del hombre,* es conciente de la tensión existente entre una explicación del comportamiento altruista que apele a la selección de grupos y una que apele a la selección individual. Allí afirma que una tribu con integrantes más fieles, cooperativos y simpáticos tendrá éxito sobre otra tribu cuyos integrantes carezcan de tales virtudes. Sin embargo, enseguida se pregunta cómo podrían tales sentimientos de fidelidad y cooperación haber emergido al interior de la tribu, pues es obvio que aquellos individuos con sentimientos de sacrificio por los demás dejarían menos descendencia que los individuos egoístas y falsos. Pág 161-2. Panamericana, Bogotá, 2002

2 *Ibid.*

entre individuos al interior del grupo, entonces los grupos con mayor porcentaje de altruistas son seleccionados. La selección de grupos es teóricamente posible, pero exige condiciones especiales y la opinión mayoritaria es que las condiciones no se dan a menudo.[3] De ahí que sea metodológicamente preferible recurrir a una explicación que se apoye en la selección individual. La selección individual se ha empleado exitosamente en la explicación de rasgos altruistas, como por ejemplo, los propios de las trabajadoras estériles en los insectos sociales. Las trabajadoras estériles ayudan a su madre a procrear en lugar de procrear ellas mismas, pero así reproducen sus genes con mayor éxito.[4] De esta manera se pueden explicar los comportamientos altruistas dentro del núcleo familiar, así como la cooperación con individuos con un grado cercano de parentesco. Sin embargo, esta estrategia no sirve para explicar la cooperación entre organismos no emparentados. ¿Cómo se explica la cooperación aparentemente desinteresada, por ejemplo entre seres humanos no relacionados por ningún tipo de parentesco?[5]

En su famoso artículo *La evolución de la cooperación*, Robert Axelrod (junto con William Hamilton),[6] intenta responder a la pregunta:

---

3 Ver por ejemplo: G.C Williams, *Adaptation and Natural Selection*, Princeton U.P. 1966; R. Dawkins, *The Selfish Gene*, Oxford U.P. 1989; John Maynard Smith, "The Concepts of Sociobiology", en: G. Stent (ed.) *Morality as a Biological Phenomenon*. Berkeley, CA: U.P. 1978. Una posición interesante en nuestro idioma es escrita por A. Rosas, "Filosofía moral del naturalismo", en *Ideas y Valores* (106). Bogotá: Universidad Nacional.

4 Ver en: Dawkins, *op.cit*, pp. 171-181 y 317-319, y E.O. Wilson, *Sobre la naturaleza humana* FCE, 1997 Cap 7.

5 El sentido de 'egoísta' y 'altruista' que uso en esta primera parte del artículo se limita al plano biológico y no supone ninguna postura *a priori* sobre la conducta conciente o intencional de los individuos. Dichos términos tampoco implican ninguna carga valorativa o moral y son usados de manera metafórica. Agradezco al profesor Eugenio Andrade del departamento de Biología de la Universidad Nacional quien, además de mostrarme la necesidad de esta aclaración, me sugiere definir 'egoísta' como: individuo cuyas acciones refuerzan su propia autoafirmación, y 'altruista' como: individuo cuyas acciones reafirman las relaciones contextuales con el grupo. Estos dos tipos de acciones no son mutuamente excluyentes, pues no puede haber autoafirmación de un agente mas que por acciones en un contexto. Si bien esta propuesta del profesor Andrade es bastante interesante y obedece a las nuevas tendencias en teoría evolutiva, aclaro que el sentido en el que tomo los dos términos aludidos es de corte más clásico, el cual es necesario para la argumentación general del presente texto. Por otro lado, en la segunda parte del artículo, 'egoísta' y 'altruísta' sí se toman en el sentido psicológico de conductas intencionales de los humanos como agentes concientes.

6 Axelrod, Robert & Hamilton, William. "The Evolution of Cooperation", en *Science*, vol. 211, 1981.

¿Cómo puede desarrollarse el comportamiento cooperativo entre individuos egoístas que están buscando siempre su propio beneficio? Como dice Axelrod, si individuos no racionales, como las bacterias, pueden exhibir estrategias altruistas que repercuten positivamente en su aptitud, con más veras lo pueden hacer individuos y naciones. Según Axelrod, el altruismo puede emerger y fijarse en un mundo asocial de individuos egoístas, sin necesidad de apelar ni a la coerción de la autoridad ni a los mecanismos evolutivos de la selección de parientes. Para que esto ocurra, el altruismo debe exhibir características conductuales específicas (que mencionaremos más adelante), las cuales, sin embargo, no necesariamente requieren racionalidad por parte de los individuos. Axelrod y Hamilton tratan el problema con un modelo formal basado en los juegos de estrategia. Si bien desde estos modelos de la teoría de juegos se aborda la acción desde una perspectiva netamente conductual, es posible hacer una interpretación motivacional de las estrategias. Es de notar que Robert Trivers,[7] en quien Axelrod y Hamilton se inspiraron, hace un dedicado análisis del aspecto psicológico que subyace al comportamiento cooperativo (como veremos luego). Tal vez por ello, Axelrod usa términos del ámbito de la psicología para referirse a las estrategias que ejecutan jugadores no conscientes en situaciones de juego ubicadas dentro del esquema del dilema del prisionero. En un ambiente donde racionalmente parece más provechoso traicionar o defeccionar, como resulta ser el marco del dilema del prisionero, no deja de ser curioso que la estrategia más exitosa (en términos adaptativos) en la versión iterada de este juego sea una estrategia 'amable' y 'misericordiosa', y que las menos exitosas sean las 'maliciosas' y 'retrecheras'. En palabras de Richard Dawkins 'nice guys finish first'- 'los chicos amables llegan primero'.[8]

Utilizando la teoría de juegos, Robert Axelrod demostró en el referido artículo cómo la cooperación puede emerger en un grupo de individuos egoístas. Para tal fin realizó un torneo por computador e instó a teóricos de juegos a someter estrategias de comportamiento

[7] Trivers, Robert. "The Evolution of Reciprocal Altruism". Reimpreso en: Clutton-Block T.H & Paul H (ed). *Readings in Sociobiology.* San Francisco, CA. Freeman, 1978.

[8] Dawkins, *op.cit*, Cap. 12.

que considerasen promisorias para competir con otras en la versión iterada del dilema del prisionero.[9] La siguiente matriz representa una movida del juego con sus utilidades ordinales:

	JUGADOR 1	
JUGADOR 2	C (3) C (3)	D (4) C (1)
	C (1) D (4)	D (2) D (2)

Figura 1.

Después de someter a prueba 14 estrategias recibidas y comparar sus porcentajes, la estrategia ganadora fue una de las más simples: TIT FOR TAT (que significa algo así como "toma y dame"), enviada por el psicólogo y teórico de juegos Anatol Rapoport. Esta estrategia consiste en cooperar a la primera movida y luego copiar la movida previa ejecutada por el oponente. Es decir, si en la primera movida el oponente cooperó, el jugador de TIT FOR TAT cooperará a su vez. Si por el contrario, el oponente defeccionó, el jugador de TIT FOR TAT defeccionará en la segunda interacción que tenga con dicho oponente, pero volverá a cooperar en la tercera interacción, si su oponente coopera en la segunda. Así, es característico de TIT FOR TAT que sólo actúa retaliatoriamente una vez, y está dispuesta a volver a cooperar como respuesta a la cooperación de

[9] Dentro de la teoría de juegos y su pertinencia para modelar situaciones sociales en donde se presenten conflictos de intereses, el dilema del prisionero es el juego más famoso y más utilizado. Este fue primeramente formalizado como un juego por M. Flood y M. Drescher en 1950, pero fue A.W. Tucker, en Princeton, quien el mismo año le dio su ilustración estándar. Este dilema describe una posible situación en donde son ofrecidos a dos prisioneros diversos tratos y penas, dependiendo de si quieren o no colaborar con la justicia. El punto es que, teniendo un orden de preferencias, los prisioneros, de manera aislada, eligen racionalmente un curso de acción -no cooperar- que termina en un resultado que va en detrimento de sus propios intereses, pues es peor para ambas partes que el resultado que hubieran obtenido de haber decidido cooperar. Una amena y bien informada introducción al dilema del prisionero, de corte histórico-divulgativo, resulta el capitulo 3 del libro de Matt Ridley: *The Origins of Virtue*, *Human Instincts and Evolution of Cooperation*, Penguin Books, NYC, 1998.

su oponente. Axelrod divulgó los resultados y programó un segundo torneo, al cual se unieron más participantes. Con un número mucho mayor de estrategias (62) enfrentándose unas con otras, sorprendentemente TIT FOR TAT ganó de nuevo, demostrando su 'robustez'. 'Robustez' en términos axelrodianos quiere decir que es una estrategia exitosa contra una amplia gama de otras estrategias. Axelrod explica entonces el éxito de una estrategia como TIT FOR TAT basándose en cuatro propiedades: abstención de innecesario conflicto al cooperar tanto como el otro jugador lo haga, provocabilidad frente a una movida no cooperativa de su oponente, capacidad de perdonar después de responder a una provocación si el oponente cambia su actitud, y transparencia de comportamiento permitiendo que el otro jugador pueda adaptar su patrón de acción[10]. Este último punto, así como el segundo, requieren por parte del individuo de una capacidad de reconocer y recordar a su oponente. En el caso de TIT FOR TAT no es necesario que vaya más allá de dos jugadas. Este éxito se debe entonces a que se utiliza una estrategia amable (nice), provocable, misericordiosa (forgiving) y transparente (clear), y esto la hace evolutivamente estable.

Ahora, ¿qué quiere decir estrategia evolutiva estable (EEE)? Una estrategia es evolutivamente estable si una población de individuos, usando tal estrategia, no puede ser invadida por una estrategia mutante. Una EEE tiene la propiedad de que si casi todos los miembros de una población la adoptan, su éxito individual resulta mucho mayor que el de cualquier mutante individual adoptando una estrategia diferente. Este resulta ser el caso, como apunta Axelrod, de la estrategia de defección en el juego del dilema del prisionero de una sola interacción: defeccionar es una estrategia evolutiva estable.[11] Los resultados de los torneos de Axelrod mostraron que TIT FOR TAT es una EEE en la versión iterada del dilema. El éxito de TIT FOR TAT se basa igualmente en la expectativa de un número indefinido de interacciones con los mismos

[10] Axelrod, Robert. *The Evolution of Cooperation,* (Basic Books: NY. 1984), Tomado del extracto publicado en *Ethics,* Singer, Peter, ed. Oxford U.P., Oxford, 1994. Página 90.

[11] Axelrod & Hamilton, *op.cit* , página 1392.

jugadores. Esto concuerda con lo que ocurre generalmente en las interacciones sociales, donde "la sombra del futuro", para usar la expresión de Axelrod, es suficientemente larga. La expectativa de interacciones futuras los induce a cooperar con el fin de aprovechar los beneficios de la cooperación, que son precisamente negados a quien no coopera.

De lo anterior podemos extraer los dos requisitos necesarios para el florecimiento de la cooperación: debe estar basada en la reciprocidad y, la sombra del futuro, es decir, la perspectiva de futuras interacciones entre dos jugadores, debe ser lo suficientemente larga para que la reciprocidad sea estable. Dadas estas condiciones, la "(…) robustez de TIT FOR TAT dependió en los torneos de tres características: nunca fue la primera en defeccionar, fue provocable a retaliar frente a la defección de otro y perdonó luego de un sólo acto de retaliación."[12] Esto muestra además que entre individuos interesados en su propio bienestar la mejor jugada no es defeccionar sino cooperar. Axelrod interpretó sus resultados diciendo que en condiciones plausiblemente realistas, la cooperación puede emerger entre individuos egoístas, sin que medie ningún tipo de autoridad. Concluyó su estudio con estas palabras:

> Estos resultados dan un cuadro cronológico para la evolución de la cooperación. La cooperación puede empezar con grupos pequeños. Esta puede florecer con reglas que son amables, provocables y algo misericordiosas. Y una vez establecida en una población, individuos usando tales estrategias discriminativas pueden protegerse ellos mismos de cualquier invasión. El nivel global de cooperación tiende hacia arriba y no hacia abajo. En otras palabras, la maquinaria para la evolución de la cooperación tiene un ratchet. [13]

El alcance del trabajo de Axelrod y Hamilton obligó a Richard Dawkins a incluir una exposición del mismo entre las adiciones

[12] Axelrod &Hamilton, *op. cit*, página 1393. TIT FOR TAT es, como vimos arriba, una estrategia 'robusta', i.e, es exitosa en términos evolutivos con respecto a otras estrategias rivales.

[13] Axelrod, *Op. Cit.*, Singer (ed.), página 92. ('Ratchet': mecanismo o dispositivo de engrane dentado implementado en una rueda unida a una barra y que permite el movimiento en una única dirección).

que hizo a la 2da. edición de su célebre *Gen egoísta*. Allí Dawkins expone cómo TIT FOR TAT puede florecer biológicamente. TIT FOR TAT puede invadir poblaciones de individuos no cooperativos si se agrupa inicialmente en agregados locales, lo cual puede suceder con el mecanismo de la selección del altruismo en contextos de parientes cercanos. En la mayoría de las especies los individuos viven más cerca de sus parientes (hermanos, padres, primos, etc.) que de otro grupo en la población. Esto se debe a la 'viscosidad' poblacional. "Viscosidad' significa cualquier tendencia de los individuos a continuar viviendo cerca al lugar donde nacieron". Y si los individuos relacionados genéticamente tienden a compartir rasgos fenotípicos, igualmente lo hacen con rasgos conductuales heredados genéticamente, como por ejemplo, jugar TIT FOR TAT.

> "Si esto pasa, individuos TIT FOR TAT, cooperando el uno con el otro en pequeños y cómodos enclaves locales, pueden prosperar tan bien que crecen de pequeños grupos locales a más grandes grupos locales. Estos grupos locales pueden crecer tanto que se esparcen hasta otras áreas, las cuales hasta ahora habían sido dominadas, numéricamente, por individuos jugando a defeccionar siempre [...] a diferencia de TIT FOR TAT, 'siempre defeccionar' no obtiene ninguna ayuda del linaje o de la viscosidad de la población."[14]

Pero lo que aquí nos interesa es la interpretación psicológica de la estrategia denominada TIT FOR TAT. Ya hemos visto que Axelrod interpretó a TIT FOR TAT como una estrategia de agentes egoístas. Sin embargo, describió las características que le dan su superioridad estratégica con términos como "nice" o "forgiving". Dawkins señala en su exposición de Axelrod que TIT FOR TAT nunca le gana a ninguna estrategia en una confrontación. Ni siquiera le gana a las estrategias que son explotables. Lo máximo que hace es empatar con ellas.[15] Si tuviésemos que hacer una interpretación

---

[14] Dawkins, Richard, *op cit*, pp . 219-220.

[15] "TIT FOR TAT de hecho nunca 'gana' un juego. Piense al respecto y verá que esta estrategia no puede obtener más puntaje que su oponente en ningún juego particular. Pero el punto es que ella tiende a lograr cada empate con un alto puntaje". Dawkins, *Ibid*. Este es precisamente el éxito de TIT FOR TAT: sumando cada juego particular logra un puntaje general mucho mayor que cualquier otra estrategia.

psicológica de esas características conductuales, diríamos que ellas son las conductas que esperamos cuando la motivación subyacente es altruista. A Axelrod aparentemente no le interesaba hacer este tipo de inferencias. Nos volvemos por ello ahora sobre el trabajo de Robert Trivers, cuyo interés por los mecanismos psicológicos subyacentes a la cooperación hace que su trabajo sea doblemente relevante para la filosofía moral.

2. La evolución de las motivaciones morales

Trivers expone en su ya clásico texto *La evolución del altruismo reciproco*,[16] el sistema psicológico que subyace al comportamiento altruista entre los humanos. El estudio de Trivers se diferencia del de Axelrod y Hamilton en que teoriza sobre las motivaciones que soportan la actitud cooperativa. Es como si sacase a relucir directamente las características psicológicas que luego fueron convertidas en etiquetas puramente conductuales en el estudio de Axelrod. Para Trivers, los motivos genuinamente altruistas son adaptativamente más exitosos que los egoístas (así tengan un manto altruista que los cubra) y por esa razón son seleccionados. Trivers define primero una 'situación altruista' como aquella en la cual un individuo, con la opción de no cooperar, dispensa a un segundo individuo un beneficio b a un costo c tal que $b >> c$ (b es mucho mayor que c). Para Trivers, las oportunidades de selección para el comportamiento altruista son grandes si se dan tres condiciones:

> "(1) cuando hay numerosas situaciones de tal tipo en la expectativa de vida de los altruistas, (2) cuando determinado altruista actúa repetidamente con el mismo pequeño conjunto de individuos, y (3) cuando pares de altruistas están expuestos 'simétricamente' a situaciones altruistas."[17]

[16] Trivers, Robert. "The Evolution of Reciprocal Altruism". Reimpreso en: Clutton-Block T.H & Paul H (ed). *Readings in Sociobiology*. San Francisco, CA. Freeman, 1978.

[17] Trivers, *Op. cit.* pp. 192-3.

Los beneficios de la reciprocidad están aquí dados por la proporción especial costo/beneficio del acto altruista, pues el costo del altruismo es compensado con creces por la reciprocidad. Trivers dice que, debido a esta proporción y teniendo en cuenta las futuras interacciones entre dos actores, la selección favorecerá a todos aquellos que son capaces de cooperar selectivamente con los que a su vez cooperan. Los que no cooperan y los que responden con defección a la cooperación (los tramposos) serán en cambio proscritos por la selección natural (esto, claro está, siempre y cuando los altruistas puedan discriminarlos). El beneficio de engañar en determinadas ocasiones a uno u otro altruista no puede compensar la pérdida de una sociedad cooperativa perdurable.

Según Trivers, durante el Pleistoceno, e incluso antes, la especie homínida pudo haberse encontrado bajo las condiciones para la evolución del altruismo recíproco: larga expectativa de vida (suficiente 'sombra del futuro'), bajo promedio de dispersión (en términos de Dawkins, alto promedio de 'viscosidad'); vida en grupos sociales pequeños, estables y con alto grado de dependencia mutua entre sus miembros.

Trivers recalca que la discriminación de los tramposos es un elemento clave en la evolución del altruismo recíproco. Hay dos tipos de engaño o de hacer trampa, una 'descarada' y otra 'sutil'. La forma descarada consiste en nunca retribuir a una acción altruista, haciendo que el altruista asuma todos los costos de la acción sin ningún beneficio a cambio. Este tipo de tramposo, obviamente, será finalmente eliminado. La otra manera de hacer trampa es la sutil. Esta implica reciprocidad, pero en un menor grado. Es decir, el beneficiario de una acción altruista devuelve menos de lo que recibe. Así, el altruista recibe un beneficio de la acción altruista pero no tanto como sucedería en una interacción equitativa. El tramposo sutil se beneficia, a su vez, más de lo que hubiera obtenido de la práctica equitativa. Esta forma de engaño, según Trivers, puede ser adaptativa, debido a que en los humanos la difusión del altruismo es evidentemente amplia, y debido a la multiplicidad de interacciones en un grupo, así como a sus diferentes ordenes de preferencias, en donde alguien puede preferir una retribución menor a la justa en lugar de nada.

Dado este cuadro inestable, donde hacer trampa es adaptativo, la selección favorecerá un sistema psicológico que regule tanto las propias tendencias altruistas y/o tramposas, así como la detección de tales tendencias de defección en los otros: "Como la selección favorece formas sutiles de hacer trampa, ésta favorecerá a su vez habilidades más eficientes para detectar tramposos"[18]. Del mismo modo, la selección favorece un sistema con la tendencia de apreciar a otros que no están directamente relacionados con nosotros, para así formar amistades altruistas: "(...) la selección favorecerá el aprecio por aquellos que son en sí mismos altruistas"[19]. Sin embargo, como hacer trampa es adaptativo, una vez hayan evolucionado estas 'emociones positivas' -como el aprecio o la amistad-, el altruista que las posee se encontrará a merced de los tramposos que son seleccionados para tomar ventaja de ellas. Es debido a esto, afirma Trivers, que se selecciona un mecanismo protector de indignación y agresión moral, el cual contrarresta las propias tendencias altruistas de seguir ejecutando actos cooperativos sin retribución alguna, al mismo tiempo que educa al tramposo por que le causa el daño inmediato o futuro de no gozar de la cooperación por parte del individuo ofendido: "Gran parte de la agresión humana tiene matices morales. Injusticia, inequidad y ausencia de reciprocidad a menudo motivan indignación o agresión"[20]. Ahora, como el tramposo tiene que cargar con la perdida de la cooperación de su contraparte cuando ésta lo sorprende cometiendo una fechoría, entonces es ventajoso (en términos adaptativos) para el tramposo, quien quiere evitar esta pérdida, hacer una compensación y garantizar, mostrando *evidencia convincente*, que se arrepiente de su fechoría y que está dispuesto a no volver a hacer trampa. A su vez, para no perder la ventaja de la futura ayuda que le pueda prestar, será ventajoso para el ofendido perdonar al tramposo: "Parece plausible, además, que la emoción de culpa haya sido seleccionada en los humanos parcialmente en orden a motivar al tramposo a compensar su fechoría y a comportarse de manera altruista en el futuro, para así evitar la

[18] *Op. Cit.* Página 212.
[19] *Op. Cit.* Página 213.
[20] *Op. Cit.* Página 214.

ruptura de relaciones recíprocas"[21]. Sin embargo, anota Trivers, una vez que las emociones que hemos enunciado (amistad, agresión moral, culpa, aprecio, etc.) han evolucionado como elementos que regulan la acción moral, la selección presionará a favor de tramposos que puedan imitar estas emociones para poder influenciar el comportamiento de los altruistas en beneficio propio. Este tipo de individuo es para Trivers un 'tramposo sutil'. Por ello, la selección debe favorecer un mecanismo que identifique y discrimine a los tramposos sutiles: "La selección puede favorecer el desconfiar de aquellos que realizan actos altruistas sin las bases emocionales de generosidad o culpa, puesto que las acciones altruistas de tales individuos se vuelven *menos confiables* en el futuro"[22]. Esto es, los actos altruistas deben ser motivados por sentimientos genuinos de cooperación, ya que un acto altruista sin tales motivos, esto es, con intenciones egoístas, no genera reciprocidad al ser juzgado como poco confiable para sustentar una sociedad cooperativa. "… el espíritu de cálculo que motiva al tramposo sutil…puede en el futuro llevarlo a hacer trampa cuando las condiciones parezcan más ventajosas"[23]. Esto es, prudencialmente me abstengo de cooperar con un individuo que no sea motivado por razones altruistas, ya que si más adelante se encuentra en una situación ideal de ventaja, puede llegar por simple cálculo racional a defeccionar. Este es un punto en Trivers que es necesario resaltar. Además de la fijación de mecanismos que se encarguen de contrarrestar los métodos cada vez más especializados y sutiles del tramposo, Trivers añade que la selección favorecerá a quienes tengan motivaciones altruistas genuinas, pues serán ellos los elegidos, por su confiabilidad, para las sociedades cooperativas. La selección natural provee entonces un mecanismo que fija los sentimientos altruistas como base de las motivaciones para cooperar. Trivers nos dice:

> Hay amplia evidencia para soportar la noción de que los humanos responden a actos altruistas de acuerdo a su percepción de los motivos

[21] *Op. Cit.* Página 216.
[22] *Op. Cit.* Página 217.
[23] *Ibid.*

> del altruista. Ellos tienden a responder más altruistamente cuando perciben al otro actuando de manera 'genuina', esto es, realizando voluntariamente un acto cooperativo como un fin en sí mismo, sin esperar ganancia.[24]

Hasta aquí Trivers ha expuesto los mecanismos altruistas y egoístas que la selección va favoreciendo en orden a regular las interacciones sociales para ventaja evolutiva del individuo. Esto podríamos ubicarlo dentro del esquema del dilema del prisionero en la versión iterada, en donde la sombra del futuro permite una aparición selectiva de mecanismos motivacionales altruistas que vienen a neutralizar las motivaciones egoístas que saltan a la luz en las interacciones bilaterales. Y esto, como dice Trivers, permite una confianza que promueve la cooperación entre dos individuos, incluso si no se conocen el uno al otro, ya que la selección podría favorecer un mecanismo para establecer nuevas relaciones de reciprocidad: "(...) un mecanismo tal podría ser el realizar actos altruistas dirigidos a extraños, o incluso enemigos, en orden a inducir amistad"[25]. Ahora, si bien los mecanismos de regulación expuestos hasta ahora se basan en relaciones bilaterales iteradas, según Trivers, estos pueden extenderse hacia interacciones más complejas, pues "Específicamente, la selección puede favorecer el aprender de experiencias altruistas y tramposas de otros, ayudando a otros a coercionar tramposos, formando sistemas múltiples de intercambio, y formulando reglas para regular dichos intercambios en tales múltiples sistemas"[26]. La reciprocidad confiable, entonces, como motivación genuina que subyace a las interacciones bilaterales de cooperación, puede hacerse extensiva en un grupo social. Esta confianza va a venir a ser el soporte de las relaciones que existen al interior de dicho grupo.

¿A donde nos lleva este análisis de Trivers? Gracias a su modelo se puede argüir que, dadas ciertas condiciones, la selección natural favorecerá no sólo comportamientos altruistas que a la larga

[24] *Op. Cit.* Página 218.
[25] *Op. Cit.* Página 219.
[26] *Op. Cit.* Página 220.

benefician al individuo que los realiza, sino también motivaciones altruistas. El análisis formal de Axelrod demostró que al interior de un grupo de individuos egoístas puede evolucionar la tendencia (conductual) a cooperar recíprocamente. Sin embargo, los requisitos operacionales de TIT FOR TAT concuerdan con los motivos psicológicos que vendrían a sustentar un altruismo recíproco entre los humanos, como muy bien vimos que expone Trivers. Según Axelrod, las estrategias más exitosas son aquellas 'amables' y 'misericordiosas'. Estas no pretenden explotar a nadie y se contentan con un empate en la interacción. Esto significa, que no quieren sacar provecho de ninguna interacción, ni siquiera si están ante estrategias incondicionalmente altruistas como 'siempre cooperar'. Además, si defeccionan lo hacen sólo en retaliación, pero olvidan enseguida la defección anterior de su contraparte si esta coopera, reiniciando la interacción con una movida cooperativa. Estas características conductuales de un jugador TIT FOR TAT, son las que esperamos, en el ámbito humano, encontrar en las personas altruistas y así lo muestra también el análisis que Trivers hace con respecto al sistema psicológico subyacente al altruismo recíproco en los humanos. Este último -por naturaleza inestable- gana estabilidad debido a la autenticidad de los sentimientos altruistas. Como dice John Mackie en su análisis de la función de la moral,[27] la estructura del castillo de naipes (como esqueleto de las relaciones sociales) adquiere con lazos menos condicionales que los egoístas-prudenciales la solidez y estabilidad necesaria para resistir a las diferentes y cambiantes fuerzas que enfrenta.

En la búsqueda de los motivos subyacentes a los actos morales están, por un lado, la motivación egoísta y por el otro la motivación altruista. Ahora, bajo el supuesto de que la cooperación tiene una existencia estable, la motivación moral debe incluir elementos altruistas. En este artículo revisamos tal idea desde una perspectiva evolutiva y concluimos que en los análisis clásicos de Trivers y de Axelrod las evidencias hablan a favor de la evolución de motivaciones genuinamente 'amables', por ser más aptas que las

[27] Mackie, John. Ethics: Inventing Right and Wrong. London: Penguin (1977).

'maliciosas'. Para ellos, nuestras motivaciones para la acción moral son genuinamente altruistas, es decir, cooperamos con el otro tomándolo como un fin y no como un medio. Con una sociedad basada en actos genuinamente altruistas se garantiza una cohesión firme de la estructura social, lo que beneficia enormemente a cada individuo perteneciente a ella.

BIBLIOGRAFÍA

Axelrod, Robert. (1984). *The Evolution of Cooperation.* New York: Basic Books, del extracto en: Peter Singer (Ed.). (1994). *Ethics.* Oxford: Oxford University Press

Axelrod, Robert & Hamilton, William. (1981). "The Evolution of Cooperation", En: *Science*, 211: 1390-6.

Darwin, Charles. (2002). *El Origen del hombre*. Bogotá: Panamericana.

Dawkins, Richard. (1976). *The Selfish Gene.* Oxford: Oxford University Press.

Mackie, John. (1977). *Ethics. Inventing right and wrong*. London: Penguin Books.

Ridley, Matt. (1998). *The Origins of Virtue*: *Human Instincts and Evolution of Cooperation*. New York: Penguin Books.

Singer, Peter (Ed.). (1994). *Ethics.* Oxford: Oxford University Press.

Trivers, Robert. (1978). "The Evolution of Reciprocal Altruism" Reimpreso en: Clutton-Block T.H & Paul H. (ed). *Readings in Sociobiology*. San Francisco, CA.: Freeman.

Williams, G.C. (1966). *Adaptation and Natural selection*. Princeton U. P.

Wilson, E.O. (1997) S*obre la naturaleza humana*. México: FCE, 1997.