

INSTITUTO DE ESTUDIOS DE LA CIENCIA
Y LA TECNOLOGÍA



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

TESIS DOCTORAL

**ESTUDIO DE LOS FUNDAMENTOS DE
LA COOPERACIÓN EN LA NATURALEZA
HUMANA DESARROLLADOS POR LAS
CIENCIAS SOCIALES**

Autor: JOSÉ LUIS HERRANZ GUILLÉN

Director: Prof. Dr. D. SANTIAGO M. LÓPEZ GARCÍA

SALAMANCA, 20 de julio de 2015

Índice

PREFACIO	9
CAPÍTULO 1.- LA LARGA CONTROVERSIA: EGOÍSMO, ALTRUISMO Y NATURALEZA HUMANA	41
1.1.- Introducción	43
1.2.- La naturaleza humana en la controversia histórica sobre el egoísmo	45
1.2.1.- De Grecia a la Ilustración	45
1.2.2.- El egoísmo y el altruismo en los debates de la Ilustración previos al surgimiento de la economía política	58
1.2.3.- La polémica en la actualidad: la sociobiología, los neohobbesianos y los neobutlerianos	67
1.3.- Los dos problemas teóricos del comportamiento altruista	73
1.3.1.- Insertar la identidad en la psicología y en las ciencias sociales	74
1.3.2.- Motivación altruista <i>versus</i> motivación egoísta: la cooperación desde la perspectiva poblacional	77
1.4.- Conclusiones	81
CAPÍTULO 2.- EL SUJETO Y LA RACIONALIDAD A LA LUZ DE LA CIENCIA CONTEMPORÁNEA	87
2.1.- Introducción	89
2.2.- La acción humana, o la libertad desde una perspectiva científica: Spinoza y la Escuela Austríaca	96
2.3.- ¿Qué es el yo? Neurobiología de la conciencia subjetiva	103
2.3.1.- El yo es un proceso neural de alto nivel	108
2.3.2.- El funcionamiento del cerebro en situaciones de pseudo-yo, no-yo, y yoes patológicos	114
2.4.- Crítica al concepto estándar de racionalidad en las ciencias sociales	121
2.4.1.- De atributo exclusivo de la naturaleza humana a modelo de racionalidad instintiva	124
2.4.2.- El protagonismo de las emociones y “demonios” en la elección racional	133
2.4.2.1.- Tiempo y conciencia	134
2.4.2.2.- Sin innatismos no hay elección racional	136
2.4.2.3.- Racionalidad y racionalización	140
2.4.3.- Refutación de la racionalidad del egoísmo optimizador	144
2.4.4.- Un utilitarismo matizado y una función de utilidad compleja: la racionalidad de la elección sin raciocinio	148
2.5.- Conclusiones	155
CAPÍTULO 3.- ALCANCE Y LÍMITES DE LAS EXPLICACIONES BASADAS EN LOS CONCEPTOS DE COOPERACIÓN RECÍPROCA Y EGOÍSMO RACIONAL	163
3.1.- Introducción	165
3.2.- Cooperación dinámica entre egoístas racionales: del Leviatán filantrópico a la reciprocidad	173
3.2.1.- Cooperación inestable en juegos de única jugada	174
3.2.2.- Cuando se cree que la interacción es inagotable: cooperación en superjuegos	179
3.3.- <i>Homo reciprocans</i> : la evolución de la cooperación mediante la estrategia del “toma y daca”	181
3.3.1.- Reciprocidad directa: respuestas de gratitud y represalia	186
3.3.2.- Reciprocidad indirecta: el valor de la reputación y la justicia comunitaria	195

3.3.2.1.- Reputación y sistema institucional	199
3.4.- Limitaciones del enfoque de la reciprocidad racional para explicar la evolución de la cooperación	206
3.4.1.- El problema de la interdependencia/complementariedad	213
3.4.2.- El problema de la “visibilidad”	216
3.5.- Conclusiones	227
CAPÍTULO 4.- LA COREOGRAFÍA OCULTA DE LA COOPERACIÓN: CAUSAS ÚLTIMAS Y SISTEMAS DE REFUERZO DEL COMPORTAMIENTO ALTRUISTA	235
4.1.- Introducción	237
4.2.- Nacidos para cooperar	243
4.3.- Cooperación más allá de los lazos relacionales y grupos genealógicos	245
4.4.- El premio a la prosocialidad	250
4.4.1.- La endo-recompensa de una homeostasis mejorada	254
4.4.2.- La exo-recompensa derivada de los sentimientos de admiración y gratitud, y los sistemas culturales de prestigio	257
4.5.- La teoría del destino común	262
4.6.- La perspectiva poblacional	273
4.6.1.- El equilibrio evolutivo entre egoístas y altruistas	276
4.6.2.- Selección discriminatoria de fenotipos	287
4.6.2.1.- Estrategias de señalización	291
4.6.2.2.- Reconocimiento de cooperadores y explotadores	297
4.7.- Conclusiones	305
INTRODUCCIÓN A LOS CAPÍTULOS 5 Y 6.- CONEXIÓN DE VIDA PSICISMO: REPLICADORES CULTURALES E IDENTIDAD HUMANA	313
CAPÍTULO 5.- LA EVOLUCIÓN CULTURAL DEL ALTRUISMO: GRUPALISMO E IDENTIDAD	325
5.1.- Introducción	327
5.2.- Dos paradojas de la cultura, y su relación con el altruismo humano	330
5.3.- Semejanza fenotípica: la hipótesis relacional y cultural de la cooperación humana.....	337
5.4.- Adoctrinamiento cultural para la convergencia fenotípica	344
5.4.1.- Marcadores e identidad: “Not by genes alone”	348
5.4.2.- Exadaptación y señuelos culturales	353
5.4.3.- Naturaleza evolutiva de los marcadores culturales de identidad	357
5.5.- Conformismo y replicación cultural	363
5.6.- El sustrato neurológico de los memes	371
5.7.- El yo, la identidad, y el sentimiento humano de lo propio	375
5.7.1.- ¿Quién soy yo? El yo-propietario a la búsqueda de su identidad	380
5.8.- Conclusiones	387
CAPÍTULO 6.- INTERSUBJETIVIDAD, RACIONALIDAD PLURAL Y SELECCIÓN DE GRUPO	397
6.1.- Introducción	399
6.2.- El comportamiento desde la perspectiva colectiva	400
6.3.- Desmitificación del solipsismo	405
6.3.1.- «¿Quién es como yo?»	407
6.4.- La neuropsicología intersubjetiva del altruismo	409
6.4.1.- La categoría instintiva del “nosotros”	419

6.5.- Psicología evolucionada para la intersubjetividad	424
6.5.1.- La perspectiva plural	428
6.5.2.- Neuroendocrinología de la afiliación y el vínculo	430
6.6.- Marcadores fenotípicos, semejanza y <i>rapport</i>	436
6.6.1.- Homofilia: cooperación entre cerebros semejantes	438
6.6.2.- Cognición cultural y altruismo en alianzas y amistades	444
6.6.3.- Construcción cultural de rasgos de humanidad y universalización de la benevolencia, o sea, de la predisposición al altruismo general	451
6.7.- Conclusiones	461
CONCLUSIONES GENERALES	469
BIBLIOGRAFÍA	481

A mis antepasados, que me legaron altruistamente el instinto de cooperar, y a mis congéneres, que me han enseñado altruistamente la virtud de cooperar. Pero sobre todo y por encima de todo a mis padres, José Luis y Marcelina, altruistas ejemplares que me entregaron el instinto y educaron, y educan, la virtud.

Prefacio

Homo sapiens es lo que los biólogos denominan “eusocial”, es decir, que está constituido por miembros de grupos que contienen múltiples generaciones y que están dispuestos a realizar actos altruistas como parte de su división del trabajo. (Edward O. Wilson 2012: 14)

Este trabajo es el resultado de una investigación cuya intuición primaria, y la correspondiente motivación para el estudio, surgieron hace bastante tiempo. En este largo periplo la actividad investigadora ha sufrido todo tipo de vicisitudes para, acumulado el nivel de maduración propicio y hallado el tiempo necesario para la redacción, plasmarse en la presente tesis doctoral.

Dos hechos observados por mí mismo fueron el acicate inicial de este trabajo académico. El primero es constatar la polaridad en la disposición y práctica del comportamiento humano hacia una inclinación prosocial o no. Observaba y observo que así como se aprecia una tendencia generalizada hacia las conductas competitivas, e incluso antisociales, existen igualmente ámbitos de la interacción social en que surgen espontáneamente conductas cooperativas entre las personas, o entre éstas y grupos o estructuras sociales. Tal actitud y motivación no se ciñe exclusivamente a las relaciones familiares o de vinculación fuerte, aunque estas son un foco principal de la motivación progrupal para la mayoría de los individuos. Otros ámbitos de la interacción social se caracterizan por la indiferencia desidentificada, la falta de estímulos para una cooperación comprometida, e incluso una abierta hostilidad relacional. La acción colectiva en procesos sociales y la acción grupal dentro de las organizaciones se ven perjudicadas por la ausencia de cooperación, o una cooperación insuficiente. El resultado de este patrón motivacional y de comportamiento no cooperativo suele ser la ineficiencia, tanto desde el punto de vista individual como colectivo. El último informe Gallup, *State of the global workplace*, sobre la implicación de los trabajadores en sus empresas, por ejemplo, pone de manifiesto que sólo el 13% de los empleados a nivel mundial se consideran a sí mismos comprometidos con su actividad en la organización en que están empleados, siendo el 63% de esta población los que se consideran no comprometidos, y el 24% los que se definen como activamente no comprometidos (O'Boyle y Harter 2013: cap. 2). El nivel de compromiso, y por lo tanto de cooperación espontánea, se pronostica deficiente en términos generales. Las empresas y otras entidades pueden invertir en identidad, y de hecho esa es una de sus actividades. Las organizaciones tienen el propósito de mejorar el compromiso de afiliados, empleados, inversores, feligreses, clientes, accionistas, etc., con los objetivos generales de la organización (Simon 2001). En la medida que la cultura organizativa (*software of the*

mind) puede ser “programada” en los cerebros, y reforzar el sentido de identidad y pertenencia al grupo, las organizaciones dedicarán recursos para crear un ecosistema cultural diferenciado y motivador (Hofstede 1991; Schein 1987; Druskat y Wolff 2005). Se trata de incrementar la motivación intrínseca para actuar, *cæteris paribus* los motivadores extrínsecos, a favor del logro de las metas organizacionales (Sheldon y Houser-Marko 2001). Sin embargo, es obvio que los resultados son claramente mejorables a nivel agregado, y que muchos casos particulares de éxito se parecen más a los planteamientos de explotación tradicionales, incluidas las actitudes stajanovistas de auto-explotación, que a modelos cooperativos de ganar-ganar.

En otro orden de cosas, durante mis años de formación universitaria en Economía me surgieron dudas importantes por las discrepancias que yo apreciaba entre determinados enfoques y conclusiones de la ciencia y mi percepción de la realidad que esa ciencia describe. Tales discrepancias son el segundo fundamento que motiva este trabajo doctoral. Una consecuencia de la extrapolación simplista de hallazgos como el del informe Gallup lleva a muchas personas a defender una visión del mundo consistente en la falta de armonía en las relaciones sociales, cuando no de un puro escepticismo antropológico. Esta mentalidad, llevada al extremo, considera al hombre un animal individualista y peligroso para sus semejantes, salvo que se le incentive o fuerce a lo contrario. Una influyente tradición doctrinal ha generalizado tales premisas conductuales de la naturaleza humana en las ciencias sociales. Este enfoque describe una naturaleza humana egoísta y socialmente autodestructiva, la de un individuo antisocial que

... puede disfrutar de la gloria (...) en un mundo de aspecto terrible (...) Casi todo vale si se desea obtener esta gloria terrenal. Las mentiras, las manipulaciones, la falta de compasión por el prójimo, la hipocresía son argucias permitidas y también recomendadas (...) no existen objetivos colectivos que perseguir o que inciten a la acción común; el hombre forma un microcosmos completo que le lleva a estructurar el mundo a su medida. (Martín y San Emeterio 2014: 79).

¿Obedece esta descripción a la generalidad del comportamiento humano en la generalidad de las relaciones intraespecíficas, o sólo a algunas de las capacidades conductuales del etograma? La experiencia propia de la mayoría de las personas enseña que esta descripción, que es un esbozo del comportamiento depredador y explotador -de

«seres mucho más peligrosos que las fieras» (*ibid.*)-, no se corresponde con la realidad de la vida cotidiana, dado que cualquier individuo psicológicamente sano reconoce, en sí mismo y en sus semejantes, la bondad e inclinación al bien común como disposiciones naturales con valor intrínseco. Gould (2010[1977]: 293) por ello ha caracterizado al *Homo sapiens* como «un animal inteligente y bondadoso». Taylor (2002: 56) destaca, en este sentido, que las principales características de la vida humana cotidiana son el cuidado de los demás y la colaboración pacífica, no el egoísmo desenfrenado.

La observación empírica informa que el ser humano no depreda sistemáticamente a sus semejantes, salvo en circunstancias extremas como pueden ser la guerra o el asesinato. La explotación intraespecífica, por otra parte, se muestra más extendida (casos extremos serían la esclavitud o los sistemas de castas), si bien, como se analizará en este trabajo, requiere un entramado institucional que la sostenga y la refuerce, y está sometida a fuerzas y equilibrios que limitan su alcance e intensidad. No obstante, en la perspectiva inversa, tampoco es cierto que seamos altruistas sistemáticos y, aunque lo fuésemos, tal disposición se manifestaría evolutivamente disfuncional.

Por ejemplo, imaginemos la situación de un incendio dentro de una estancia que amenaza la vida de dos personas. Sólo es posible escapar por una estrecha cavidad por la cual cabe exclusivamente una de ellas. La operación es compleja, y el tiempo necesario para escapar por esa salida hace que sólo sea posible que uno de los individuos pueda salvarse de morir abrasado. Estamos pues ante un caso típico de jugada única de dilema del prisionero. Tanto si suponemos que ambos individuos son egoístas carentes de escrúpulos, como si suponemos que se trata de dos monjes altruistas cuyo objetivo es el ejercicio incondicional de la caridad, en ambos casos pugnarán, ya sea por salir o no salir, y esto conduce a una tragedia individual y colectiva porque se eleva notablemente la probabilidad de que los dos individuos perezcan. La predicción de este experimento, que podemos denominar paradoja de los monjes altruistas, coincide con la del experimento llevado a cabo con egoístas que sólo se preocupan por su supervivencia y bienestar. Y no se trata solamente de una paradoja,

es además un contrasentido evolutivo, ya que teóricamente la naturaleza no selecciona comportamientos ordinarios conducentes a callejones sin salida o paradojas del tipo perder-perder. Esto quiere decir que los organismos que sistemáticamente cayeran en paradojas de esta naturaleza, o que fueran incapaces de aprovechar eficientemente sus interacciones, deberían haberse ya extinguido.

Desde este punto de vista, la aparente discrepancia entre egoísmo y altruismo no es tal ya que, al fin y a la postre, tanto los egoístas sin escrúpulos como los monjes se están comportando de hecho maximizando una función de utilidad individual. O sea, psicológicamente todos actúan de hecho como egoístas, dado que ya sea su objetivo el bienestar propio o el ajeno, son inflexibles en una concepción individualista de su lugar en el mundo que no considera objetivos ni resultados concernientes más allá del sujeto.

En este experimento mental, que quitándole su carga de dramatismo puede resultar representativo de muchas situaciones cotidianas, se parte de una noción atomista de la identidad, y esa concepción de la identidad es lo que explica la tragedia consecuente. Como explica Schmid (2005b), el juego así planteado no es en realidad un dilema, sino una paradoja práctica, según la cual los agentes se hallan encasillados en una identidad tan reduccionista que de hecho están forzados a elegir la alternativa más perniciosa. De este modo, ya maximice el agente la utilidad propia o la ajena, en ambos casos el individuo se comporta egoístamente. El altruismo, así pues, no se refiere al sujeto en particular hacia el que se enfocan las preferencias del agente, sino a la consideración de la naturaleza de su identidad plural, qué población la comparte, y qué consecuencias tiene esto en la perspectiva desde la cual elige el individuo cuando se enfrenta a un dilema social que puede requerir, o no, su sacrificio personal pero que, en cualquier caso, será en pos de objetivos de supervivencia referidos a una población y no únicamente a él.

La noción de identidad no es pues un aspecto irrelevante en la interacción social de una especie como la humana¹. Las especies sociales se caracterizan porque practican

¹ Actualmente diversos estudios han enfatizado el factor identidad, el pensamiento social y los modelos mentales de pensamiento, como aspectos explicativos de la idiosincrasia y del éxito diferencial de los modelos de desarrollo social, institucional y económico. Véanse, p.ej. *World Development Report 2015*, del Banco Mundial, y Salom y Farinós (2014).

asiduamente estrategias de cooperación para sobrevivir. Sin embargo, es preciso destacar que aunque los individuos *necesiten* cooperar para el logro de *sus* objetivos, ello no es un argumento suficiente para que se produzca la cooperación, y mucho menos para que evolucione. Por ejemplo, explica Gauthier (1994[1986]: 35) que: «Sólo los seres cuyas aptitudes físicas y mentales son en general iguales o mutuamente complementarias pueden esperar que la cooperación resulte beneficiosa para todos». La naturaleza ha seleccionado diversas maneras de favorecer el encuentro entre individuos dispuestos a cooperar en pos de logro de objetivos mutuamente beneficiosos, y una de ellas, la genuinamente humana, es el dispositivo neuropsicológico que representa a los individuos dentro de una identidad en la que se comparten metas. Los objetivos son comunes en la medida que existe un sustrato común de supervivencia: un *pool* de replicadores diseminado en la población de individuos.

Efectivamente, la teoría evolutiva comparte con la economía que las adaptaciones que conforman el etograma humano (algo así como la función de utilidad de la especie) es el resultado de un presunto cálculo y balance de costes y beneficios, lo cual, dicho sea de paso, concuerda con la noción estándar de comportamiento racional (aun siendo este altruista) (Khalil 2004: 104). Sin embargo, para la biología evolutiva tal balance obedece a equilibrios poblacionales en los que el protagonista no es el organismo-individuo, sino el grupo de rasgos al servicio del que existe dicho organismo, y gracias al cual puede sobrevivir y replicarse. En la teoría evolutiva las fuerzas de selección actúan sobre linajes de rasgos homogéneos, y no sobre individuos particulares (Dugatkin 2007: 25). Dentro de esos linajes poblacionales se pueden admitir grupos emparentados biológica o culturalmente, como las familias genealógicas o las identidades culturales. Evidentemente, en este segundo caso la especie ha tenido que desarrollar la base de la cultura, que son los símbolos, hasta convertirse en una especie simbólica (Broncano 2004). De este modo, el patrón de comportamiento adaptado desde la perspectiva de los replicadores, y no del organismo, es *ayuda a los que son como tú y te estarás ayudando a ti mismo*.

Esta investigación es un esfuerzo para profundizar y esclarecer en las controversias acerca de la naturaleza del altruismo humano, un fenómeno que, como

explica Gordon (1995: 23), se trata de uno de los principales objetos de interés en la literatura del pensamiento social, pero de difícil estudio, y al que todavía no se ha conseguido integrar en un modelo analítico general del comportamiento humano. Los problemas para abordar el estudio del altruismo humano surgieron ya en los primeros intentos científicos. Darwin, por ejemplo, encontró dificultades para encajar el altruismo en su modelo general de evolución por selección individual, lo que le condujo a expresar una de las primeras formulaciones de la selección de grupo (Darwin 1972[1871]: I, v). En *La ayuda mutua* (1902) Kropotkin sostuvo que los darwinistas de su época habían descuidado estudiar el papel desempeñado por la cooperación entre organismos de la misma especie como estrategia adaptativa de supervivencia. En realidad, hasta la década de 1960 la cooperación y el altruismo no formaron parte de la agenda de la investigación biológica de vanguardia, y sólo hace un par de décadas la cooperación ha ido entrando, con una relevancia creciente y generalizada, en los programas de investigación de las ciencias sociales y las neurociencias. Edward O. Wilson caracterizó el comportamiento altruista como «el problema teórico central» de la sociobiología (Wilson 1980[1975]: 3), una opinión que 25 años después se mantiene en las palabras de Lord Robert May, quien en su discurso presidencial de apertura de la Conferencia Anual de la Royal Society expuso en 2005 que:

La pregunta más importante todavía no contestada en la biología evolutiva, y por extensión en las ciencias sociales, es cómo ha evolucionado el comportamiento cooperativo y puede mantenerse en los grupos y sociedades de humanos u otros animales. (Citado en Rockenbach 2007: 215).

Recientemente Tomasello, *et. al.* (2012: 673) se han referido al altruismo como la gran incógnita evolutiva de la naturaleza humana. En efecto, se trata de un tema de estudio con larga tradición intelectual y académica pero que, como apostilla Taylor (2002: 196), cualquier estudioso actual es conocedor de que todavía no contamos con todas las piezas que permiten formular una teoría general de amplio alcance. Podemos decir que a pesar de ser un contenido clásico en la indagación del ser humano sobre sí mismo, el altruismo es un tema de estudio de actualidad, sobre el que aún no existe amplio consenso científico, y que en muchas ocasiones adolece de atávicos defectos de enfoque, entre la indefinición y la ambigüedad (Khalil 2004).

El estudio de la cooperación ha adquirido una significativa importancia en las últimas décadas en la biología, psicología social, teoría de juegos y la economía, así como otras ciencias que estudian el comportamiento social humano. La evolución de la cooperación humana fue elegida en el año 2005, en el número especial conmemorativo del 125º aniversario de la revista *Science*, uno de los 25 problemas científicos más importantes por resolver en el s. XXI en una encuesta realizada a la comunidad científica internacional (Penisi 2005).

En los encuentros científicos del ámbito de la genética y la evolución biológica celebrados en la última década se ha concluido que la cooperación basada en la reciprocidad es un fenómeno bastante difundido en la naturaleza, y es el principal factor explicativo de comportamientos animales que pueden calificarse de proto-cooperativos, desde los microorganismos hasta los animales de inteligencia más desarrollada². Martin Nowak (2006a) defiende la tesis, ya planteada anteriormente por los biólogos Maynard Smith y Szathmary (1999), de que la evolución es un proceso constructivo debido a la cooperación: la cooperación es un fenómeno imprescindible para explicar cómo la vida evoluciona construyendo nuevos niveles de organización, desde los protozoos unicelulares hasta los animales multicelulares, y desde el individuo, pasando por el pequeño grupo, hasta las sociedades multitudinarias y mucho más complejas. En este sentido la cooperación es una forma de generar nuevas combinaciones innovadoras impulsadas por el motor de la selección natural, como anticipara la teoría de la endosimbiosis formulada por Lynn Margulis en la década de 1970.

La cooperación ha sido estudiada desde una amplia diversidad de perspectivas y modelos, tanto en animales no humanos como en la especie humana. Quedan muchos interrogantes por resolver, especialmente teniendo en cuenta que una de las características más misteriosas de la cooperación humana es que los seres humanos a veces cooperan anónimamente con personas desconocidas y con las que seguramente no volverán a repetir interacción, ni ellos ni sus descendientes o amigos. Es lo que en teoría de juegos se denomina cooperar en interacción anónima de única jugada. En economía

² Véase Hammerstein (2003), que contiene las ponencias de la 90ª Conferencia Dahlem sobre “Genetic and Cultural Evolution of Cooperation” celebrada en Berlín en junio de 2002; así como las principales conclusiones que relatan Sigmund y Nowak (2000) acerca del Encuentro Internacional “The Evolution of Cooperation and Communication” celebrado en Steyr (Austria) en julio de 2000.

la temática de la cooperación en la acción colectiva también ha experimentado un destacable impulso dentro de los últimos avances acaecidos en su rama experimental y en la neuroeconomía, y además ha recibido el galardón del premio Nobel concedido en 2009 a Elinor Ostrom, estudiosa de la organización social que dedicó buena parte de su trabajo científico a estudiar las instituciones de la cooperación humana, y a refutar la hipotética tragedia de los bienes comunales.

Centrar el contenido del tema de estudio no es fácil. Ello se debe a que a nuestro juicio el tratamiento académico dado a la cooperación se caracteriza por ser “caleidoscópico”, esto es, no abordado directa sino indirectamente, y desde una diversidad de enfoques parciales, con frecuencia imprecisos y a veces cargados de connotaciones y juicios preanalíticos. Sin embargo, en la actualidad se cuenta con suficiente trabajo empírico y teórico como para integrar lo que puede aunarse y desechar lo que no tiene justificación. Así pues, este trabajo viene a ser un tamiz, una clasificación, una bitácora, un catalizador que puede resultar útil para la formulación futura de una teoría general.

El altruismo se enmarca en la gama de los llamados comportamientos prosociales, los cuales incluyen una amplia gama de conductas con el denominador común de la ayuda y la cooperación (Penner, *et. al.* 2005). Se trata de acciones en las que el agente, pudiendo hacerlo, sin embargo no sigue la estrategia de optimización individual. La característica de los comportamientos altruistas es la acción de transferir recursos sin expectativa de reciprocidad. En psicología la motivación es un aspecto relevante de la acción, por lo que el carácter prosocial o no de la conducta no se ciñe exclusivamente a la transferencia, sino que resulta relacionalmente significativa la intención de ayudar como motivo raíz de la acción. Al contrario, la acción de transferir puede enmascarar la intención primaria de ayudarse a sí mismo a través de una ayuda a otra persona (p.ej. para mejorar la reputación social, aliviar el sentimiento de culpa, o el estrés psíquico causado por la observación de una persona en dificultades) (Batson 1991). En economía y biología, no obstante, la mayoría de los modelos no requieren que la motivación sea relevante a la hora de explicar el carácter prosocial de los comportamientos. En ambas disciplinas se considera prosocial la conducta altruista, y tal comportamiento consiste simplemente en actuar soportando un coste en favor del

beneficio de otro individuo.

Raimo Tuomela indica que la cooperación es un concepto complejo y difícil de manejar, y explica que la aproximación más adecuada a su estudio consiste no en analizarla desde una metodología *a priori*, sino en investigar empíricamente las condiciones y circunstancias que la hacen posible (Tuomela 2000: 1-2). Entre los economistas actuales interesados en explicar la cooperación humana con modelos matemáticos, como es el caso de Herbert Gintis, Samuel Bowles y sus colaboradores, se recomienda abordar el estudio del fenómeno cooperativo desde una investigación transdisciplinar y desde enfoques que precisan de una constante innovación metodológica (Gintis 2007a).

Desde la perspectiva de las ciencias naturales, la biología evolucionista moderna define los actos de cooperación como comportamientos benéficos para otros individuos, pero que conllevan un coste para el individuo que los realiza (Dugatkin 2007: 61-62). Para la mayoría de los biólogos y antropólogos evolucionistas esa definición de cooperación coincide exactamente con la de altruismo, es decir, se trata de dos términos sinónimos (p.ej. Silk y Boyd 2010: 224)³. Sin embargo, Dugatkin precisa que todavía no se ha alcanzado un amplio consenso académico acerca de la definición exacta de altruismo en biología. La teoría de juegos (que es una disciplina académica independiente, pero también una herramienta de análisis utilizada tanto en biología evolutiva como en economía) suele utilizar con más frecuencia el término cooperación que el de altruismo, y en sus supuestos metodológicos asume un significado para la cooperación similar al de la biología evolutiva⁴.

Otros enfoques se centran en mayor medida en presentar la cooperación desde el punto de vista de la coordinación, premeditada o no, de acciones entre diversos individuos (Colman 2005), lo cual no necesariamente implica una pérdida para alguno en pos de una ganancia para otro/s, sino que es posible una coordinación con resultados individuales de suma positiva en el logro conjunto de una meta. La perspectiva

³ Por el mismo razonamiento puede definirse el egoísmo como el comportamiento consistente en conseguir un beneficio a costa del perjuicio de otros (Magee 2006: 257).

⁴ Sober y Wilson (2000: 65) resaltan que la teoría de juegos utiliza el término cooperación y raramente el de altruismo porque en realidad el altruismo no se contempla como supuesto motivacional por los teóricos de juegos.

sociológica, por ejemplo de Nisbet (1974: 182), enfatiza el ingrediente relacional de la cooperación, entendida como colaboración en orden al logro de una meta común, ya sea un interés común o la esperanza de una recompensa. Una dirección parecida se encuentra en la perspectiva biológica que plantea Ruse (1994: 287 y ss.), según la cual no toda cooperación ha de perjudicar los intereses reproductivos de los individuos que la practican. Esta misma vendría a ser la aproximación teórica desde la economía, que suele separar claramente la cooperación del altruismo y, en lo concerniente al altruismo, no siempre es considerado como el trasvase unilateral neto de recursos *inales* (en los que caben, por ejemplo, experiencias utilitarias como la seguridad personal, el orgullo, el amor, la alegría o simplemente el bienestar psicológico) de un individuo a otro. En economía el agente siempre maximiza utilidad, sea cual sea la experiencia o proceso orgánico al que se denomine de tal modo, y por ello todo comportamiento altruista es en realidad la gestión de una ganancia utilitaria individual supeditada a la transferencia unidireccional de recursos⁵. No obstante, la economía, a diferencia de los enfoques funcionalistas de la cooperación y del altruismo predominantes en sociología, se decanta por explicar cómo es posible que surja la cooperación, concepto económico más amplio que el de altruismo, entre individuos que no persiguen metas comunes, ni comparten replicadores o identidad⁶.

La perspectiva más extendida es la del altruismo definido como entrega no condicionada de recursos. El primatólogo Frans De Waal, con una visión descriptiva propia de la etología tradicional, incide en la idea del altruismo como ayuda, consistente en el sacrificio por los demás (De Waal 1997: 22). Una concepción también centrada en el sacrificio prosocial es la de Toro Ibáñez y Castro Nogueira (1999: 43), que lo explican como transferencia neta de eficacia biológica. El filósofo Philip Kitcher (2006:

⁵ Véanse, entre otros, los trabajos siguientes de Becker: (1997[1974]; 1997[1976a]; 1987; 1981). Una perspectiva general de las investigaciones sobre el altruismo y la filantropía en la economía puede encontrarse en Fontaine (2000; y 2007).

⁶ Téngase en cuenta que los conceptos de cooperación y altruismo, que ya aparecen en algunos referentes principales de la Historia del Pensamiento Económico, sin embargo, han recibido un tratamiento académico formal relativamente reciente en las instituciones gestoras del conocimiento. Así, la palabra “cooperación” no existe como tal actualmente en el *JEL Classification System for Journal Articles* de la American Economic Association. Sí existe el código C71, “Juegos Cooperativos”, introducida en 1991. Asimismo, el código “altruismo” es el D64, y fue introducido en 1993 dentro de un epígrafe especializado en temas de economía del bienestar. (Agradezco esta información a Drucilla Ekwurzel, Directora del Servicio de Publicaciones, Marketing y Producción de EconLit, de la American Economic Association).

127) se refiere al altruismo en términos de generosidad voluntaria hacia otro individuo sin expectativa de recompensa, ya se trate de una acción consciente y premeditada o no. Elliott Sober (1994: 154), por su parte, considera que el altruismo es un comportamiento perjudicial para el individuo que lo ejercita, pero ventajoso para el grupo al que pertenece.

En el *Diccionario de la Evolución* de Richard Milner (1995) no aparece la entrada “cooperación”, aunque sí existe la de “altruismo” (p. 39), y llama la atención que cuando se acude a esta última no figura ninguna definición o texto explicativo, sino que el autor remite directamente a las entradas “gen egoísta” (p. 291) y “selección de parentesco” (p. 572). El altruismo es visto por economistas y biólogos como una forma de egoísmo, mediante cuya práctica el agente obtiene algún beneficio a corto plazo (utilidad) o a largo (propagación de sus genes). En esta retórica encaja la frase de Michael Ghiselin en su obra *The economy of nature and the evolution of sex* (1974) cuando desacredita la hipótesis de un altruismo genuino en las relaciones humanas, afirmando irónicamente que si se araña a un (presunto) altruista, se puede comprobar como sangra un hipócrita (real) (*op. cit.*, p. 247). O sea, el altruismo viene a ser la excepción que cumple la regla del egoísmo.

Ghiselin sostiene que altruismo y cooperación son conceptos difícilmente manejables por la ciencia, dada su carga de prejuicios ideológicos y morales. El autor incluso avanza aún más y niega teóricamente la realidad del altruismo. Sin embargo, otras aproximaciones se decantan por la apreciación opuesta: que la vida social empírica es un tejido de actos altruistas (Eccles 1996: 114). En esta línea afirmativa, numerosos estudios acreditan que con frecuencia e intensidad nada desdeñables la gente corriente comparte el conocimiento sin contraprestación y, en general, ayuda a desconocidos en juegos de única tirada. Las personas también participan en huelgas y manifestaciones (a veces exponiéndose a graves peligros), votan en procesos electorales, dan propinas y limosnas, realizan donaciones a colectas benéficas, participan en la gestión del bien común a través de las instituciones políticas y religiosas. Igualmente se ayuda en accidentes y catástrofes. La gente en ocasiones se expone al oprobio de los demás expresando abiertamente sus puntos de vista, participa en actividades de voluntariado, y presta su colaboración a personas enfermas o debilitadas. Con frecuencia estos actos de

altruismo se realizan en el anonimato, y los beneficiarios son personas desconocidas con quienes se presume no se volverá a interactuar. Estas son muestras cotidianas de generosidad no necesariamente heroicas. En otras ocasiones, no obstante, la generosidad adquiere tintes heroicos, e incluso dramáticos, como las acciones de entregar la vida directamente por los demás en los conflictos bélicos y las catástrofes, u otras con menor carga de dramatismo y no por ello menos heroicas, como es el caso de la donación en vida de órganos y tejidos, la actividad de los disidentes en regímenes políticos totalitarios, o la de los misioneros en sociedades que persiguen su religión, y toda la tipología de héroes anónimos en el ejercicio profesional de la medicina, la seguridad ciudadana o la asistencia social. ¿Qué racionalidad puede haber en que un individuo sacrifique su propia vida por la de otra u otras personas?

En un intento de clarificar conceptualmente y sistematizar el tratamiento y la clasificación de los comportamientos prosociales, Herbert Simon (1983: 57-58) ha planteado una distinción elemental, tomada del mirmecólogo Edward O. Wilson (1980: 158-159), entre comportamientos de cooperación fuerte y débil⁷. La primera, la cooperación fuerte o incondicional (*strong altruism* en terminología de Simon), se refiere a los comportamientos humanos voluntarios que benefician a otros individuos sin expectativa alguna de recibir una recompensa o intento de evitar un castigo. El segundo tipo de cooperación alude a la reciprocidad o cooperación condicionada (*weak altruism*, o *enlightened self-interest*, en palabras del autor), de manera que el cooperador espera recibir en algún momento (directa o indirectamente) algún tipo de recompensa, o bien evitar algún modo de castigo. En este sentido la cooperación débil puede ser una estrategia de comportamiento capaz de evolucionar, si bien expuesta a riesgos asociados a la posible interacción con explotadores y depredadores. La cooperación fuerte, en su caso, resulta *a priori* manifiestamente inepta para la supervivencia evolutiva, aparenta ser una anomalía por tratarse de un comportamiento económicamente “irracional”. La cooperación débil podría tener una lógica de supervivencia apta si el agente es capaz de estimar correctamente sus expectativas sobre la reciprocidad de la contraparte. No obstante, tanto la una como la otra son estrategias de comportamiento que en principio parten de una desventaja relativa en comparación con otras estrategias no cooperativas.

⁷ Simon en realidad utiliza el término “altruismo” en lugar de cooperación, siguiendo la conceptualización de Wilson, quien denomina a ambos mecanismos de ayuda altruismo de “núcleo fuerte” y de “núcleo blando” respectivamente.

Dentro de la cooperación fuerte podríamos incluir aquellos actos de altruismo real que surgen de acciones impulsivas, como por ejemplo las causadas por las emociones, ansias y arrebatos (Elster 2001). Igualmente el caso de otros procesos psíquicos irreflexivos conducentes a acciones arriesgadas en determinada tipología de individuos (Taylor 2002: cap. 9). Sin embargo, no se puede generalizar que toda acción de cooperación fuerte, o altruismo estricto, sea un comportamiento impulsivo, inconsciente o imprudente, o algo así como una “tontería sentimental” (*sentimental foolishness*) carente de recursos o sentido racionales (Khalil 2004: 109).

El modelo explicativo del altruismo que ha conseguido una mayor aceptación científica es la teoría de la evolución del altruismo por eficacia inclusiva de William Hamilton (1963; y 1964)⁸. El modelo de Hamilton parte de la premisa de que el comportamiento animal es un producto de la evolución codificado genéticamente, igual que la morfo-fisiología del individuo en cuestión. Toda codificación genética, se supone, está adaptada para sobrevivir y replicarse “viajando” en organismos que los genes contribuyen a crear, y que son vehículos eficientes de los mismos⁹. El altruismo es un tipo de comportamiento codificado por la genética, igual que el egoísmo, y puede definirse, siguiendo a Dawkins (1993[1976]: 5), como la conducta que contribuye a aumentar las oportunidades de supervivencia de otro organismo a expensas de las propias.

El axioma de la replicación fue establecido por Darwin en el campo de la ciencia, quien afirmó, en *El origen de las especies*, que todos los organismos se esfuerzan naturalmente hasta el extremo por aumentar en número, sin que se conozca

⁸ La formulación original de la hipótesis se debe, no obstante, a J. B. S. Haldane, quien la presentó en un trabajo científico publicado en 1955 sobre genética de poblaciones (Haldane 1955). Años antes la había bosquejado en un libro divulgativo dirigido al gran público (Haldane 1932), que hoy día es considerado un clásico en biología. Aquí exponemos el modelo hamiltoniano de eficacia inclusiva que figura en Dugatkin (2007: cap. 5).

⁹ La biología evolutiva de la síntesis moderna teoriza que la unidad básica de análisis es el gen, y no el organismo, de modo que son los genes las entidades que se replican, no los organismos, y son las responsables de configurar organismos portadores (vehículos) de sí mismos. La unidad de selección es un conjunto coherente e integrado de genes que codifican un organismo y su conducta. La interacción de los genes que integran el ADN es la causa de un organismo que, interaccionando con otros organismos y con el medio, consigue -o no- replicar (y no en su totalidad, en los organismos que se reproducen sexualmente) los genes que componen su ADN, y que por lo tanto lo configuran como organismo. Sin embargo, los procesos de selección (natural y sexual, social y cultural) actúan sobre fenotipos (o sea, sobre organismos), no sobre genotipos ni genes específicos, o al menos no lo hacen de una manera directa (Simon 1983: 49).

cuáles son las causas de tal predisposición universal (Darwin 1998a[1859]: 90). El replicador es la unidad básica de herencia en la biología, y coincide con genes o grupos de genes que expresan rasgos fenotípicos en un organismo y que pasan de generación en generación. Un replicador es por tanto una entidad que cumple dos propiedades en lo que aquí nos afecta (Binmore 2009: 188; Pinker 2004: 208): (a) se replica a sí misma produciendo copias con rasgos idénticos o bastante parecidos, y capacidades también para replicarse; y (b) constituye un patrón determinado de comportamiento estratégico¹⁰. En teoría todo replicador está determinado a propagarse de tal modo que, de no encontrar restricciones, llenaría el universo de copias de sí mismo. La finitud de los recursos y la coexistencia de replicadores conduce a la competencia entre ellos, que se valen de “vehículos” eficientes para viajar en el espacio y en el tiempo. Esos vehículos de la biosfera son los organismos¹¹.

La clave del modelo de Hamilton es que contiene un parámetro de calibración del comportamiento altruista en función de la semejanza existente en cuanto a replicadores que portan los respectivos vehículos orgánicos. Esta semejanza interpreta la identidad de los individuos en función de los genes que portan y que determinan su comportamiento, de ahí que la eficacia que persiguen los organismos no sea la del propio individuo, sino la de aquella población de individuos que comparten una identidad genética¹². Así, siendo c el coste que soporta el individuo altruista por su

¹⁰ La caracterización expuesta nos parece adecuada para nuestro propósito, sin embargo, es preciso indicar que el concepto de replicador no está teóricamente cerrado, y por tanto todavía es objeto de debate: por un lado, la replicación genética neo-darwiniana *versus* la replicación cultural neolamarckista (Dawkins 1990; Blackmore 1999). Por otro lado, las controversias internas dentro de la biología acerca de la unidad básica de análisis (gen, ADN, individuo, grupo) (Sober 1994: cap. 4). Y en tercer lugar, las controversias relacionadas con la propia caracterización de una entidad replicadora. Así, por ejemplo, Hull (1988: 408) define un replicador como unidad que transmite su estructura idéntica en sucesivas repeticiones. Por su parte Godfrey-Smith (2000: 414) no hace hincapié en el grado de exactitud de la estructura que se transmite, sino que se queda en la semejanza genérica y resalta la causalidad de esa semejanza.

¹¹ Una concepción holista del concepto de vehículo apuntaría a mecanismos de soporte de los replicadores más amplios y abstractos, como el grupo, la sociedad, e incluso el entramado de normas sociales y morales (véanse Dawkins 1990, y Ruse 1994: cap. 6). Asimismo, un concepto más holista del replicador conduciría a unidades de análisis más amplias o sistémicas, como el cerebro individual, la interacción entre cerebros, el grupo, e incluso las redes exosomáticas de la sociedad cultural al completo (véanse Fischer 1990, y Bartra 2006).

¹² Utilizando palabras de Steven Pinker:

Un gen que actuara replicando copias de sí mismo en el interior de las gónadas de algún *otro* animal sería tan bueno como un gen que las replicara en el interior de las gónadas *del propio* animal. En la medida que se trata del gen, una copia vale lo mismo que otra copia, y es irrelevante el animal que haga las veces de anfitrión. Para un gen encargado de la construcción del cerebro, lo único especial acerca de las gónadas del animal es la *certeza* de que las copias del gen se hallarán en esas gónadas. (Pinker 2004: 513; cursivas en el original).

acción de ayuda a otro individuo, y siendo b la ayuda específica que recibe el beneficiario del acto altruista, el altruismo será una conducta eficiente siempre que se cumpla la siguiente relación entre las variables c y b ¹³:

$$r \cdot b > c ,$$

donde r (coeficiente de relación genética) es la probabilidad de que el individuo receptor de la ayuda disponga en su ADN del mismo alelo que codifica el patrón de comportamiento altruista en el individuo que le proporciona la ayuda, y tanto c como b están medidas en términos de eficacia (*fitness*) perdida por el donante y ganada por el beneficiario de la ayuda¹⁴.

Introducir la variable r en este modelo tiene importantes ventajas a efectos explicativos. La primera es que r es una variable cuantificable, dado que es posible medir la correspondencia genética entre dos organismos¹⁵. La segunda es que al tratarse de una probabilidad, su valor necesariamente será tal que $0 \leq r \leq 1$. Conviene resaltar que en este modelo la cooperación (entendida como altruismo) puede explicarse sin necesidad de que exista reciprocidad (directa o indirecta). Lo relevante aquí es si el organismo que dispone del gen que codifica el comportamiento de cooperación es capaz de responder eficazmente a los estímulos que lo impelen a sacrificarse por (o tener en cuenta las necesidades de) sus semejantes.

En este modelo de eficacia inclusiva el éxito en el logro del objetivo perseguido no se mide en términos del individuo, sino en términos de balances poblacionales: el

¹³ Aquí se entiende por eficiencia que el comportamiento del individuo favorezca la replicación del gen que codifica su comportamiento altruista. O sea, que pueda evolucionar. Si se optimiza o no, es un aspecto que cabe ser considerado en el modelo.

¹⁴ La expresión formal de la regla de Hamilton puede ampliarse para explicar el acto de ayuda altruista a más de un individuo simultáneamente. En estos casos la cooperación tendrá sentido evolutivo si el acto de ayuda conlleva un beneficio total superior al costo que tal acto supone para el altruista. La fórmula quedaría como sigue: $\sum_{(i, 1, N)} [r \cdot b] > c$, donde N es la cantidad total de beneficiarios de la ayuda.

¹⁵ La fórmula general que cuantifica el valor del coeficiente de relación parental entre los individuos A y B es:

$$r_{AB} = (1/2)^g ,$$

siendo g la distancia generacional existente entre ambos, que se calcula como la suma de generaciones entre el individuo A y el antepasado común que tenga con el individuo B, más las que separan al antepasado común del individuo B.

peso relativo del grupo de organismos que disponen de un alelo específico (el que codifica el comportamiento altruista) en el conjunto de la población. De este modo, los individuos capacitados genéticamente para el altruismo se comportarán con mayor generosidad cooperativa con aquellos otros individuos a los que se vinculen mediante una relación r más elevada. Asimismo, dado un r^* fijo, el individuo capacitado para el altruismo será tanto más altruista cuanto menor sea el cociente (c/b) asociado a su relación con el individuo beneficiario. Esto último explica que incluso con un coeficiente de relación genética muy bajo (cercano a cero) puede ser eficiente el altruismo de pequeños favores a desconocidos, e incluso también puede serlo un muy costoso altruismo hacia un desconocido si el donante puede asegurarse que b es muy elevado respecto a c (es decir, si el “precio” relativo de la ayuda es relativamente bajo, lo cual situaría el fenómeno del altruismo en el plano de los incentivos)¹⁶. En cualquier caso, el tipo de preferencias que está planteando Hamilton en su modelo llevarían a un supuesto agente racional altruista a preferir: (1) cooperar preferiblemente con otros individuos que tienen una elevada probabilidad de ser altruistas como él (afines genéticos), frente a hacerlo con extraños de los que no se conoce verosímelmente su identidad; (2) cooperar con aquellos que *cæteris paribus* están más necesitados de ayuda en términos relativos (es decir, los que la unidad marginal de ayuda -al coste marginal c' - les repercute en mayor beneficio marginal b'); y (3) en circunstancias de incertidumbre sobre el nivel de afinidad genética y sobre el nivel de necesidad de los posibles receptores, y disponiendo de una cantidad fija de recursos para cooperar, preferir cooperar con más individuos que con menos.

Hamilton refiere su teoría al altruismo dentro de la red de individuos que comparten una identidad genética, y concretamente a los parientes consanguíneos. Es por esto que Edward O. Wilson (1998: 162) utilizó la frase “Blood is thicker than water” para ilustrar la visión hamiltoniana del altruismo. Los intentos posteriores de teorización del altruismo han aprovechado la sobriedad del modelo de Hamilton, inspirándose en él, si bien, y especialmente desde el punto de vista de las ciencias sociales, se ha intentado ampliar la versión hamiltoniana a una identidad no sólo genética.

¹⁶ En este aspecto, un modelo demostrativo de la generalización de la regla de Hamilton a individuos no emparentados es el de Alger y Weibull (2012).

Algunos modelos más recientes que se han construido sobre la idea de Hamilton de que existe una lógica, racionalmente explicable, subyacente al comportamiento altruista. En ellos, sin embargo, se reconocen implícita o explícitamente diversos puntos débiles que existen en el planteamiento hamiltoniano. A enumerar, entre otros:

- a) En primer lugar, y lo que supone la debilidad más importante, el modelo se enfrenta a problemas de verificación. La teoría de la selección de parentesco se ha demostrado incorrecta matemática y biológicamente (Nowak, *et. al.* 2010).
- b) Por otra parte, no se ha llegado a reconocer la causalidad estrictamente genética del comportamiento altruista. O sea, no es posible determinar la existencia de un gen, o entramado de genes, responsable del comportamiento altruista.
- c) Además, es evidente que existe una significativa cantidad de altruismo exofamiliar.
- d) La teoría de Hamilton es una teoría biológica presentada para la generalidad del fenómeno vida. Sin embargo no profundiza en la experiencia psicológica del comportamiento altruista propia de la especie *Homo sapiens*. Es decir, falta una psicología del altruismo, o la integración del altruismo psicológico en el comportamiento de la especie.
- e) Otro punto a destacar es que en la teoría no se contempla que el comportamiento altruista pueda ser adaptativo no sólo en el sentido de la eficacia de la población de genes hacia el que se ejecuta, sino también para los genes concretos que constituyen el organismo altruista.
- f) Puesto que la selección natural opera directamente sobre fenotipos, y sólo indirectamente sobre los genotipos, pueden existir otros replicadores explicativos del fenotipo (p.ej. los replicadores culturales de la hipótesis meme) cuya consideración en el modelo contribuyan a mejorar su capacidad explicativa del comportamiento altruista empírico.

g) Y para finalizar, el modelo de selección de parentesco puede sustituirse por modelos estándar de genética de poblaciones aplicados a niveles múltiples de la selección natural. Se trata de modelos de selección natural en varios niveles, o multinivel, que posee los dos elementos necesarios para la evolución: por un lado, se ha observado que existen rasgos a nivel grupal (que incluyen la capacidad de cooperar, la empatía y la organización en redes) que son heredables en los humanos; por lo tanto, tales rasgos varían genéticamente en cierto nivel de una persona a otra. En segundo lugar, la cooperación, la cohesión y la unidad de propósito afectan de modo manifiesto a la supervivencia de los grupos diferenciados (que no necesariamente han de competir o estar en conflicto entre ellos), y por ende a los individuos que los integran (Wilson 2012: 336).

Uno de estos modelos recientes es la teoría de la cooperación formulada por los economistas Bowles y Gintis (2008) en *The New Palgrave*, consistente en una variante de la fórmula de Hamilton en la que el individuo altruista no introduce el filtro de semejanza genética para discriminar hacia qué otro individuo dirige su acción de ayuda en pos de una cooperación eficiente. Los autores inician el artículo monográfico afirmando que existe cooperación cuando al menos dos individuos actúan comprometiéndose en acciones conjuntas de las que extraen beneficios mutuos. O sea, identifican la cooperación con formas de simbiosis o mutualismo, y para ilustrar el fenómeno citan algunos ejemplos. Entre estos mencionan el intercambio de bienes, el pago de impuestos para financiar bienes públicos, la producción en equipo, la gestión de un conjunto de recursos comunes, la colusión entre empresas, el acto de votar a favor de programas de redistribución de la renta en beneficio de otros, la participación en acciones colectivas tales como una manifestación, y el cumplir normas socialmente benéficas¹⁷. Esta definición introductoria no especifica la estructura ni el flujo de costes y beneficios asociados con la acción, por lo que no se estaría en disposición de evaluar el nivel de eficiencia asociado al comportamiento. Sin embargo, en el mismo artículo

¹⁷ Otros ejemplos descriptivos de lo que es cooperación citados en la literatura reciente son: compartir alimentos (y recursos en general) para asegurarse frente a las épocas de carestía, arriesgar la vida en la guerra para proteger al grupo, trabajar conjuntamente en la construcción de canales y fortificaciones, y castigar a los asesinos y ladrones para mantener el orden social (Boyd y Mathew 2007: 1.858). También: cumplir las reglas sociales y los compromisos asumidos, comportarse lealmente, mantener sus promesas y cumplir sus amenazas, incluso cuando tales acciones vayan claramente en contra del propio interés (Nesse 2001: 12). Y asimismo: participar en huelgas, asistir a protestas políticas, o participar en boicots a productos (Kanazawa 2004a: 45).

los autores ofrecen una explicación teórica de la cooperación, y coincidente con la expresada por ellos mismos en un trabajo previo. Consiste en el comportamiento de incurrir en costes personales para participar en una acción conjunta la cual aporta beneficios superiores a esos costes a otros miembros del grupo (Bowles y Gintis 2003: 429-430)¹⁸.

La explicación formal de esta segunda definición vendría a ser como sigue. Siendo $c > 0$ el coste que un individuo soporta personalmente por proporcionar un beneficio b a otro individuo, el primer individuo estará cooperando con el segundo si se cumple que $(b - c) > 0$. Se trata de una proposición con la que se quiere explicar una cooperación eficiente, en la que los agentes se comportan como si tendieran hacia un equilibrio estándar en el que el ingreso marginal (Img) tiende a igualarse al coste marginal (Cmg) en cada uno de los individuos¹⁹. Aquí, no obstante, caben dos posibilidades si conectamos la teoría con la división planteada por Simon antes explicada. Así,

- a) Hipótesis del *individualismo metodológico*, o *selección individual* (altruismo débil, reciprocidad o cooperación condicionada, en terminología de Simon): esa tendencia hacia el equilibrio microeconómico requiere que el individuo receptor del beneficio b (el individuo 2) se comporte recíprocamente, y en ese mismo momento o en otro momento futuro aporte al que le ayudó (al individuo 1) un beneficio b' , soportando un coste c' , de tal modo que $(b' - c') > 0$; y además $Img_1(b')$ tienda a igualarse en el equilibrio al $Cmg_1(c)$, e $Img_2(b)$ tienda a igualarse en el equilibrio al $Cmg_2(c')$. Los individuos cooperarán si existe una elevada probabilidad de reciprocidad y mientras que $Img > Cmg$. Como se puede comprender, aquí el beneficio relevante a efectos de optimización es el del agente. En esencia, esta explicación es la del *altruismo recíproco* que introdujo el biólogo George Williams (1996[1966]), y que ha sido desarrollada y

¹⁸ Obsérvese la apostilla de que la acción cooperativa está dirigida *a los miembros del grupo* al que pertenece el cooperador. Una matización idéntica es la que hace Sober (1994: 154), cuando se refiere al altruismo como un acto que perjudica al individuo que lo realiza, pero que es «ventajoso para el grupo».

¹⁹ No especificamos la unidad de medida de estos ingresos y costes marginales. Puede ser utilidad, posibilidades de supervivencia individual, posibilidades de replicación genética, valoración de flujos energéticos, etc. En cualquier caso, se trata de valoraciones marginales para cada uno de los individuos (o conjuntos de individuos).

divulgada por el también biólogo Robert Trivers (1978[1971]). El que Williams y Trivers denominaran “altruismo” a esta conducta resulta quizás inexacto y puede dar lugar a confusión, dado que en estas interacciones los individuos pueden cooperar motivados exclusivamente por su propio interés. Quizás una denominación más adecuada pudiera ser *cooperación por turnos*. Esta hipótesis de la cooperación, recuperada por Axelrod y Hamilton (1981) y estudiada ampliamente por Axelrod (véase bibliografía), incluye la que utiliza la teoría económica y otras disciplinas que se basan en la teoría de la elección racional²⁰.

- b) Otra posibilidad es admitir la hipótesis del *grupalismo metodológico* o *selección grupal* (altruismo fuerte, activo o incondicional, en el lenguaje de Simon). En este patrón de comportamiento el individuo es capaz de cooperar dentro de un grupo (al que pueden pertenecer o no parientes consanguíneos), y por ello mejora sus posibilidades de replicación individual respecto a otras situaciones alternativas, como por ejemplo no cooperar con nadie y vivir en la autarquía completa, o ser un explotador interno en grupos bien capacitados para luchar contra el oportunismo. A diferencia del individualismo conductual, en este patrón de selección de grupo el beneficio relevante a efectos de optimización no es el del agente, sino el de la población. El problema que surge ahora es explicar cómo la evolución ha generado adaptaciones que hagan compatible estos comportamientos progrupales con la supervivencia evolutiva del propio altruismo. Hoy día está admitido por la neurociencia cognitiva y la psicología social que los individuos humanos son capaces de identificarse cognitiva y/o emocionalmente con otros individuos, empatizar y simpatizar con ellos, y de ahí la motivación para realizar comportamientos cooperativos de *altruismo activo o fuerte* (Tyler y Blader 2001; Decety y Chaminade 2003a y 2003b; Vignemont y Singer 2006; Singer y Steinbeis 2009). Este altruismo consiste en soportar el coste c para que otro individuo obtenga un beneficio b , de tal modo que $(b - c) >$

²⁰ En efecto, la motivación de todo intercambio comercial en el sentido convencional de la teoría económica es que el valor de b' para el individuo que comercia sea superior a c , o sea que $(b' > c)$. Por el contrario, en la motivación puramente cooperativa el intercambio recíproco tiene como objetivo primario que $(b > c)$. La diferencia es importante: en el primer caso el individuo 1 persigue su propio interés buscando una ganancia neta *para él*; en el segundo se persigue el interés del otro, si bien sujeto a la restricción de que exista una ganancia neta para el conjunto $(b - c > 0)$. Esta ganancia, una vez ejercida la reciprocidad por parte de 2, se ve incrementada todavía más, puesto que ambas partes ganan aun no siendo esa la intención de cada individuo por separado.

0; pero, a diferencia de la cooperación dentro del parámetro del individualismo metodológico, aquí no es necesaria la reciprocidad para que la interacción tienda al equilibrio²¹. En el altruismo fuerte los individuos actúan como si tendieran hacia un equilibrio en el que $Img_2(b) = Cmg_1(c)$, y esto nos llevaría a que el cooperador se comportara como si $Img_1(b) = Img_2(b) = Cmg_1(c)$. En esencia esta explicación es el modelo hamiltoniano con el supuesto de que $r^* = 1$, y desligado de la consanguinidad; o sea, nos salimos de la eficacia inclusiva teórica convencional. Se puede decir que en las conductas de altruismo fuerte el cooperador concibe al individuo beneficiario de su cooperación como si fuera él mismo, así se trataría de una “buena copia” o un desdoblamiento *psicológico*²². Obviamente, al tratarse esa buena copia de una concepción mental-representacional, no asociada *a priori* con la identidad genética, la eficacia inclusiva de la que estaríamos tratando aquí es una eficacia inclusiva *ampliada*, o *eficacia poblacional múltiple*, que puede integrar más organismos que los emparentados por consanguinidad en la familia genealógica directa. Esta hipótesis requiere utilizar un enfoque de psicología evolutiva que explique este tipo de preferencias, pues niega el supuesto económico del propio interés, y manifiesta la incapacidad para explicar el altruismo de la tesis biológica neodarwinista basada únicamente en la selección (genética) individual. Si no se introduce la selección multinivel no se comprende aquí el altruismo. Los individuos han de ser altruistas dentro de las estructuras relacionales de su grupo, o sea, dentro de su población, para que esta pueda competir con otras poblaciones por el aprovechamiento de los recursos y el logro de sus objetivos (Traulsen y Nowak 2006; Sober y Wilson 2000).

¿Cómo es posible que el agente benefactor del grupo, o incluso de la especie, cuyo comportamiento no está motivado por su propio interés *qua* organismo, sino por lo que podemos llamar intereses generales o bien común, pueda haber evolucionado frente a otras estrategias diferentes e interaccionando con ellas en un ambiente de recursos

²¹ Herbert Gintis (2006b) concreta el concepto de altruismo (en su sentido de máximos) explicando que el acto altruista beneficia a otro individuo con un coste para el altruista, de modo que no es posible un beneficio para él ni siquiera en el largo plazo. O sea, a largo plazo alguien se beneficia de una externalidad positiva no internalizable.

²² Estas experiencias son posibles por adaptaciones neuropsicológicas con valor de supervivencia, desarrolladas en la filogénesis de la especie (Lamm, *et. al.* 2007).

escasos? Existen tres posibles explicaciones: (a) si admitimos que el propósito o intención consciente es un aspecto relevante en la interacción social humana, la primera es que el cooperador reciba beneficios, de forma directa o indirecta, de la red social en que opera o de su propia configuración orgánica, sin que estos sean el propósito de su acción benefactora. (b) La segunda es que el cooperador esté optimizando la asignación de recursos en orden a maximizar la eficacia integrada del *pool* de replicadores al que pertenece. Esta población puede ser un *pool* de rasgos homogéneos de tipo genético o bien cultural. En este segundo caso se estaría beneficiando a otras copias de sí mismo, que son patrones homogéneos de acción, de modo que la probabilidad de supervivencia de la población resulte ampliada. En un análisis evolutivo, todo comportamiento viene impulsado por motivaciones y reforzado por experiencias psicológicas bien adaptadas para el propósito de la eficacia, ya sea esta hamiltoniana o multinivel, de aquellos replicadores en interacción, biológicos y culturales. Y (c), incluso un individuo no-altruista puede llevar a cabo acciones cooperativas dentro de su población, ya sean producto de automatismos semiconscientes o de procesos deliberativos, en primer lugar por conveniencia, para poder sobrevivir y replicarse en ambientes caracterizados por la interacción de grupos que compiten entre ellos, en los que la cooperación interna fortalece su capacidad competitiva y expansiva como población. En segundo lugar, el egoísta -al igual que el altruista- se puede comportar altruistamente dentro de su grupo o población para aumentar de forma indirecta la eficacia inclusiva de la población de afines diseminada entre diversas poblaciones (integradas por altruistas y egoístas mezclados) que compiten entre ellas.

¿Resulta posible, así pues, integrar las dos vías de cooperación que establece Simon en una única vía explicativa de la cooperación general? Aquí planteamos que sí es posible tal integración, puesto que las conductas prosociales pueden definirse como aquellas en que el individuo actúa en favor de la eficacia del grupo al que pertenece. Estas conductas pueden reforzar su eficacia como organismo singular (selección individual), a la vez que, o bien la de otros organismos con los que se comparte la pertenencia a una población de replicadores (selección grupal). El resultado es compatible con la evolución dentro de procesos de selección en varios niveles, o selección multinivel (Sober y Wilson 2000).

En esta investigación defendemos la tesis de que cooperar no es una elección o una preferencia utilitaria, al estilo de como se plantea en economía la función de utilidad altruista: un utilitarismo *à la* Ghiselin. Antes bien, cooperar consiste en la expresión de un rasgo epigenético seleccionado en la evolución de *Homo sapiens*, que está latente en el etograma, y que se expresa a través de procesos mentales adaptados para las formas de vida social características de la especie. Esto es posible porque los humanos disponen de un sistema coherente y estable de evaluaciones cognitivas y emocionales en la interacción social que los economistas denominan “función de utilidad” (Robson 2002; Cosmides y Tooby 2005). Ahora bien, hablamos de una función de utilidad diferente a la que se muestra en las escuelas de economía, pues implica: (a) la captación de, e identificación con, las utilidades de otros, o al menos una simulación de las mismas, en la motivación del comportamiento individual. Y (b) la existencia en el etograma de preferencias grupales muy básicas en favor de la cooperación, cuyos equilibrios cambian debido a la dinámica interna de la población, que contribuye en mayor o menor medida a reforzar, o no, esta dotación del etograma individual.

Utilizando a la inversa la famosa metáfora de la mano invisible formulada por Adam Smith (quien creía que el comportamiento humano es pro-específico), podríamos hablar de una estructura universal de función de utilidad, de naturaleza cooperativa, que motiva a los individuos en pos de la supervivencia de agregados supra-individuales. Este algoritmo psicológico se sustenta en un sistema neurológico modular (cognitivo, emocional y operativo) que capacita para la sociabilidad a través de la cooperación. Se trata de una adaptación que actúa previamente al razonamiento, al cálculo utilitario y a la elección consciente, y conduce al individuo por un “espejismo” en el cual se esfuerza por la supervivencia poblacional, aunque generalmente él no lo sabe y no es capaz de formularlo (Burnham y Johnson 2005; Rosenberg y Linquist 2005; Wilson 2007; Bergstrom 2002).

En una alegórica relación principal – agente, en la que el principal es la Naturaleza y el agente el individuo, puede formularse que el principal establece una meta final y un sistema de incentivos para evitar los clásicos problemas de agencia. El objetivo final es la expansión de la vida, y el sistema de incentivos viene a ser lograrlo

a través de la maximización de la eficacia biológica de cada individuo (Robson 2002: 91). Esto conlleva la interacción competitiva entre individuos, y también entre las diferentes formas de vida. Sin embargo, la maximización de la eficacia individual incluye asignar altruistamente recursos para la supervivencia entre los individuos con que el agente puede establecer relaciones de cooperación eficientes. Este agregado de interdependencias actuales o potenciales constituye una población.

Así, cuando los humanos cooperamos entre nosotros y actuamos favoreciendo los intereses de otros individuos, nos comportamos como si fuésemos guiados por una mano invisible que, conduciéndonos a ayudar a los que son semejantes a nosotros (es decir, a aquellos que en el caso actuarían de igual manera), permite que nos beneficiemos de ventajas individuales que no figuraban en el propósito de la cooperación. Con frecuencia esas ventajas repercuten en los individuos que cooperan directa y discretamente, y otras veces los beneficiarios no son los individuos *qua* sujetos, sino otros que portan rasgos semejantes, y que por lo tanto pertenecen a su linaje biológico o cultural. Esto es un beneficio de tipo continuo, que afecta a grupos y redes de individuos concebibles metodológicamente como un flujo o población.

Una conclusión de esto consiste en que las poblaciones pueden considerarse, a efectos metodológicos, como si se tratara de individuos compuestos, con intereses coordinados, que evolucionan con su propia ordenación de preferencias, así como sus propias “creencias” y su propia “racionalidad” de grupo (Herrmann-Pillath 1994). La función de utilidad poblacional es un sistema psicológico de toma de decisiones que se conforma como consecuencia de la interacción entre la genética y el ambiente. En ella opera un cálculo económico supra-individual. Esta disposición cooperativa esencial y universal de la función de utilidad, en los individuos no aquejados de sociopatías, es el producto de otra “mano invisible” u orden espontáneo que hoy, en el s. XXI, podemos denominar *evolución* (Cosmides y Tooby 1994b)²³.

²³ La inclusión del concepto de función de utilidad en la perspectiva evolucionista permite contemplar la utilidad como una variable continua. El sesgo estático de la mayoría de los modelos microeconómicos confiere a la utilidad un carácter discreto coherente con el individualismo metodológico, pero chocante con una visión evolucionista de la economía y de la naturaleza humana, que aquí vamos a utilizar como herramienta de análisis de la cooperación.

La presente tesis doctoral es un estudio acerca de los fundamentos que sobre la cooperación en la naturaleza humana han sido desarrollados por las ciencias sociales a la luz de los nuevos métodos, conceptos y descubrimientos surgidos en ese mismo campo, y en otras disciplinas científicas tales como, principalmente, la biología y la psicología evolutivas, y las neurociencias. El estudio se estructura en los siguientes capítulos:

En el capítulo 1 se sintetiza el carácter de la controversia secular sobre el egoísmo y el altruismo en la naturaleza humana, una controversia de origen filosófico que tiene su reflejo en las actuales ciencias del comportamiento. Se parte del pensamiento en la antigua Grecia para, pasando por el hito filosófico de la Ilustración, llegar hasta la revisión de la polémica en la actualidad, así como los principales problemas teóricos que se plantean en torno al comportamiento altruista: (a) insertar el concepto de identidad en la psicología y en las ciencias sociales, y (b) situar la cooperación dentro de la perspectiva poblacional. De este capítulo surgen tres derivaciones que son tratadas en los siguientes capítulos: (1ª) el sujeto; (2ª) la racionalidad, y (3ª) la relación del sujeto con sus semejantes, o sea, por un lado el fenómeno de la identidad, y, por otro, su naturaleza relacional: la intersubjetividad.

En el capítulo 2 se estudia la naturaleza científica del sujeto y de la noción de racionalidad, dos aspectos de importancia singular en la teorización del comportamiento altruista dentro de unas coordenadas analíticas propias del s. XXI.

En el capítulo 3 se presenta la teoría de la cooperación construida sobre los supuestos del egoísmo racional y la reciprocidad, así como el alcance y los límites de estas concepciones y sus implicaciones en la capacidad explicativa de los modelos de evolución de la cooperación.

En el capítulo 4 se estudian las causas últimas y los sistemas de refuerzo adaptados que hacen posible la evolución del comportamiento altruista en la naturaleza humana. Se analiza la importancia que tiene en el estudio del altruismo los estudios sobre la presencia de innatismos adaptados para reforzar la eficacia del cooperador a

través de endo y exo recompensas, así como la teoría del destino común en cuanto teoría social de la cooperación. Asimismo se expone la perspectiva poblacional de la cooperación, en la que incluimos las implicaciones que han tenido los estudios que introducen el equilibrio evolutivo entre egoístas y altruistas, y las adaptaciones de estrategia pro-cooperativas en orden a posibilitar una selección discriminatoria de fenotipos.

En los capítulos 5 y 6 se aborda la cooperación desde la perspectiva del psiquismo humano, lo cual requiere una introducción previa sobre los replicadores culturales y la identidad humana, esto es, la cultura y los modelos de selección cultural. Nos adentramos en el planteamiento de la evolución de la cooperación por selección grupal/multinivel. Así, en el capítulo 5 se trata específicamente la evolución cultural del altruismo a través de los fenómenos del grupalismo y la identidad, representativos de las fuerzas de selección grupal que operan dentro de la replicación de rasgos fenotípicos. Tal es el caso del adoctrinamiento cultural, y la discriminación efectiva en favor de los marcadores de identidad, o discriminación a favor de la semejanza fenotípica. En este capítulo también se presenta una teoría psicológica evolutiva del sujeto fundamentada en la identidad y en la noción humana de propiedad, es decir, el yo explicado como un nudo de propiedades e incumbencias.

En el capítulo 6 se estudian las aportaciones de la neuropsicología de la intersubjetividad, con sus implicaciones en el concepto de racionalidad plural, y el resultado de la evolución de la cooperación por selección de grupo en poblaciones de rasgos epigenéticos semejantes.

Son muchas las personas e instituciones que han contribuido con su ayuda u omisión de ayuda a que esta investigación haya avanzado y finalizado. No es preciso enumerarlas aquí a todas. Sin embargo, sí desearía mencionar expresamente la gratitud debida a quienes en los principios, hace ya bastante tiempo, cooperaron conmigo siendo yo un joven e inexperto profesor universitario de Economía. Este es el caso de Antonio

Soto, funcionario de la Organización Internacional del Trabajo (OIT), así como el propio departamento de Cooperación de dicha organización, en cuanto que me ayudaron en todo aspecto durante mi estancia de investigación en Ginebra (Suiza). Asimismo deseo expresar mi gratitud a Aline Pawlowska y Bruce Thordarson, directivos de la Alianza Cooperativa Internacional (ACI), y a Rubén Villa y Miguel Ángel Verdura, de la por entonces Unión de Cooperativas Madrileñas de Trabajo Asociado (UCMTA). La Escuela de Estudios Cooperativos de Facultad de Ciencias Económicas y Empresariales de la Universidad Complutense de Madrid, y especialmente quien por entonces era su director, el profesor Carlos García-Gutiérrez Fernández, me ayudaron en los primeros pasos, encauzándome hacia una orientación que después fue abandonada, pero que constituyó para mi una experiencia valiosa. Igualmente podría decir de los profesores Victoriano Martín Martín, Manuel Santos Redondo, y Carlos Rodríguez Braun, del departamento de Historia e Instituciones Económicas I de la Facultad de Ciencias Económicas y Empresariales de la Universidad Complutense de Madrid, quienes me ayudaron a esclarecer definitivamente cuáles no debían ser los derroteros de mi vocación y esfuerzo investigador. La profesora Estrella Trincado Aznar, de ese mismo departamento, leyó con paciencia y dedicación algunos borradores iniciales del proyecto de tesis, aportando, igual que la filósofa Mónica Cavallé Cruz, diversos consejos y dudas importantes sobre el propósito, formulación y alcance de esta investigación.

Mi paso por los departamentos de Historia e Instituciones Económicas de la Universidad Complutense de Madrid, Universidad Europea de Madrid, y Universidad Carlos III, constituyó una experiencia vital intelectualmente revitalizadora que nunca olvidaré, como tampoco la posterior experiencia y aprendizaje adquiridos en el departamento de Economía de la Empresa de la facultad de Economía de la Universidad Carlos III. Capítulo diferente fue mi experiencia profesional de dirección de la revista profesional *Sociedad Cooperativa. Revista de Economía Social*, publicación especializada en economía social y cooperativa perteneciente al grupo editorial WoltersKluwer, donde tuve la oportunidad de relacionarme con directivos y trabajadores del tejido empresarial y organizativo de las entidades del llamado “tercer sector”.

Trabajar en el Ayuntamiento de Getafe, administración pública donde estoy ocupado profesionalmente como funcionario de carrera desde el año 2002, me ha permitido ganar el sustento con que sobrevivir mientras investigaba a saltos y con no pocas tribulaciones. Salvo contadas e importantes excepciones, en esta organización he podido conocer directamente en qué consiste la cultura de la no cooperación, lo cual ha enriquecido empíricamente las bases analíticas y ha estimulado mi motivación para completar este trabajo académico.

Para finalizar, mis buenos amigos carmelitas descalzos del Centro Internacional Teresiano Sanjuanista (CiTeS), en Ávila, y especialmente su director, Francisco Javier Sancho Fermín, y mi actual esposa, Eloína García Galdós, aportaron a mi trabajo la claridad y el ejemplo de lo que consiste el espíritu fraterno de cooperación y comunidad, además de la visión de tres grandes figuras del pensamiento cristiano cuyas doctrinas no son ajenas a los principios sustentadores del altruismo en la cultura universal: Teresa de Jesús, Juan de la Cruz, y Edith Stein.

Una mención especial y aparte merecen dos personas por su influencia decisiva en el devenir de esta investigación. La primera es la profesora Rocío Orsi Portalo, de la Universidad Carlos III, recientemente fallecida siendo muy joven aún, quien me ayudó a comprender la presencia seminal en la filosofía del debate secular que enfrenta a dos visiones de la naturaleza humana, y a la postre dos posiciones existenciales: la egoísta-individualista, y la cooperativa-comunitarista. La segunda es el director de esta tesis doctoral, el profesor Santiago M. López García, de la Universidad de Salamanca, quien ha desempeñado un protagonismo capital sin el que, a parte de la retórica al uso, este trabajo doctoral realmente no hubiera sido posible. Santiago López ha sido una inspiración para mí desde que ocasionalmente leí su artículo, escrito junto al profesor Jesús M^a Valdaliso, de la Universidad de País Vasco, “Economía, biología y evolución. Algunas reflexiones sobre la 'Economía evolutiva' y la importancia de la historia”, publicado en la revista *Anthropos* en el año 1999. Durante la última década Santiago ha sido un exigente y controvertido, a la par que sumamente generoso y creativo director de mi trabajo y aún más, un maestro que me ayudó a introducir el darwinismo y la evolución en mi modo de comprender la realidad de la naturaleza humana. Obviamente,

su huella en esta tesis doctoral es lo suficientemente profunda como para eclipsar muchas otras influencias. Mis limitaciones y torpezas, sin embargo, salen a la luz a pesar del exigente trabajo de dirección realizado por el profesor López.

Capítulo 1.- La larga controversia: Egoísmo, altruismo y naturaleza humana

Lupus est homo homini, non homo, quom qualis sit non novit.
(Lobo es el hombre para el hombre, y no hombre, cuando desconoce quién es el otro). (Plauto, *Asinaria. La comedia de los asnos*, 495).

Iustitiae debetur quod homo homini sit deus non lupus.
(A la justicia es debido que el hombre sea un dios para el hombre y no un lobo). (Francis Bacon. *Novum Organum*, §129).

... diariamente la experiencia (muestra) con tantos y tan impresionantes testimonios que está prácticamente en boca de todos el dicho: “el hombre es un dios para el hombre”. (Baruch Spinoza, *Ética*, IV, xxxv).

El hombre es un animal que no sólo está intensamente interesado en su propia preservación, sino que también posee un fino y primigenio sentido acerca de su valor como hombre. Apartarse de esto es causa de no menos perjuicios corporales y materiales. En el nombre del hombre mismo yace ya una cierta idea de dignidad, de modo que la respuesta más efectiva a la insolencia y a los insultos es decir: “Mira; no soy un perro, sino un hombre como tú”. (Samuel Pufendorf, citado en Padgen 2001: 30-31).

Das una limosna y me preguntas qué te reporta eso a ti. Mi respuesta más certera es: “*Has contribuido a que el destino de otro sea mucho más llevadero y nada más*”. Si esta respuesta no te ayuda ni tampoco te importa, entonces no has querido dar realmente una limosna, sino hacer una transacción económica, y ahora sientes que has sido timado, lo cual debe atribuirse a tu falta de entendimiento. Si, por el contrario, te importa que *aquel* que soporta carencias sufra menos, entonces has alcanzado tu objetivo y has contribuido a que sufra menos, y ves exactamente cómo tu dádiva se recompensa con creces. (Schopenhauer (2011[1822]): 134).

1.1.- Introducción

Un debate clásico ha enfrentado a dos corrientes de pensamiento que se remontan a la antigüedad, y que de alguna manera se manifiestan en la filosofía de todas las épocas históricas posteriores a través de dos concepciones alternativas de la naturaleza humana: la egoísta-utilitarista y la cooperativa o prosocial. La primera sostiene que el ser humano es un animal egoísta, que sólo se ocupa de sus intereses individuales, y que concibe a los demás como objetos de los que extraer alguna ganancia personal. Los supuestos del egoísmo son un axioma para esta corriente, es decir, el hombre no puede ser de otra manera. La segunda tradición no niega el supuesto del egoísmo, pero se abre a motivaciones que incluyen los intereses y necesidades de otros individuos como fines en sí mismos, y no sólo como medios instrumentales para obtener algún beneficio.

Ambas tradiciones de pensamiento se han ido reproduciendo en renovadas polémicas que han llegado a la actualidad. Entender las principales coordenadas intelectuales de esta controversia es uno de los objetivos del presente capítulo. No se puede obviar que el estudio científico de la cooperación, que ya es posible hoy día, parte de los conceptos, ejes argumentales e ideas centrales que nos han llegado procedentes de las controversias filosóficas del pasado.

Desde un enfoque de estudios sociales de la ciencia y la tecnología estamos en uno de esos momentos clave en que los instrumentos de análisis y los modelos conceptuales permiten una exploración netamente científica, pero a la vez la exploración ha de partir de los supuestos filosóficos que hemos estado utilizando hasta hoy mismo. ¿Cuánto de lo que la filosofía planteó está siendo utilizado como hipótesis de partida en el análisis científico (neurológico, por ejemplo) del comportamiento altruista?

Fueron pocos los pensadores que intuyeron el funcionamiento de la mente y cómo ese funcionamiento explica el altruismo. Su pensamiento se presentará en los siguientes capítulos a éste. Sin embargo, no se pueden pasar por alto las ideas filosóficas contrapuestas a la hora de entender el altruismo. Conocer esta polémica sirve para entender los antecedentes y atavismos con que se plantean las hipótesis sobre el altruismo a la luz de los hallazgos de las neurociencias.

En el apartado 2 de este capítulo trazamos un boceto de los principales argumentos y puntos de confrontación entre ambas visiones de la naturaleza humana. Como comprobaremos, es inevitable empezar por la filosofía de la Grecia clásica, puesto que Platón fija la condición egoísta como un principio inherente a la naturaleza humana. Continuaremos con los hitos más destacados de la filosofía en Roma y el cristianismo medieval. Seguidamente nos centramos en la transformación que experimentaron los debates con las renovadas ideas materialistas y hedonistas del Renacimiento, que contribuyeron a alimentar nuevas versiones de la polémica en la época de la Ilustración (apartado 2.2), especialmente sobre la naturaleza de las pasiones y de la psicología humana.

En la actualidad la polémica ha pasado del campo de estudio filosófico al científico, y concretamente se enmarca de manera destacada en la sociobiología, y lo que se ha dado en llamar posturas neohobbesiana y neobutleriana en las ciencias sociales del comportamiento humano, tal y como se explica en el apartado 2.3. Una vez planteado el panorama general hasta la actualidad, en el apartado 3 se acomete una exposición sobre los dos puntos más controvertidos a la hora de analizar el comportamiento altruista:

- a) El carácter de la motivación humana y la capacidad (o no) del individuo de actuar motivado por los intereses y necesidades de otros individuos. Y como consecuencia de lo anterior,
- b) la perspectiva poblacional que se ha de aplicar en el estudio del comportamiento altruista. Para estudiar el comportamiento egoísta sólo hace falta un individuo, sin embargo, para el estudio del altruismo se ha de comprender al individuo dentro de un equilibrio poblacional al que se adaptará o intentará modificar.

Es en estas fronteras donde aún tienen vigencia las hipótesis de corte filosófico a la vez que van creciendo y aumentando su capacidad explicativa las aportaciones de la biología y las neurociencias. Por tanto, antes de enfrentarlas a los conceptos e hipótesis de las neurociencias y la biología, es preciso conocer hasta dónde se ha llegado y cuáles son las principales hipótesis filosóficas sobre el egoísmo y el altruismo humanos.

1.2.- La naturaleza humana en la controversia histórica sobre el egoísmo

La humanidad de lo humano ha sido un tema de agitada disputa filosófica a lo largo de la historia. En definitiva, de lo que se trata es de si es posible reconocer algún tipo de característica representativa que posibilite tratar igualmente lo que es igual, y así aprovechar las posibilidades que ofrece la cooperación entre los individuos, para así evitar las desastrosas consecuencias sociales del egoísmo cortoplacista y las disputas violentas. Se trata de una innovación cultural compleja (metafísico-religiosa) que apuntala los posibles déficit cognitivos, y emocionales, a la hora de categorizar aquellos “objetos” que gozan de la dignidad privilegiada de humanos, y que no tienen, por ejemplo, los demás animales y vegetales que utilizamos (y matamos) para sobrevivir. Hoy día la ciencia define la naturaleza humana por su genoma, y no por la morfofisiología, los atributos mentales o los comportamientos observados. En el pasado, no obstante, la condición humana siempre fue objeto de polémicas en torno a estos aspectos. Dentro del pensamiento occidental tales polémicas pueden remontarse a Grecia y llegan hasta la llamada *Polémica de los naturales*.

1.2.1.- De Grecia a la Ilustración

En *La República* (II, 2, 2; 359b-359d) Platón desarrolla una reflexión acerca de la tipología de los deseos a colación de la justicia, que es el *leitmotiv* de toda la obra²⁴. Platón sostiene que existen bienes que son deseados por sí mismos, así como otros que se valoran sólo por su utilidad y que son deseados de manera instrumental. El comportamiento justo y virtuoso es uno de estos últimos, ya que se trata de «bienes penosos que no merecen nuestros cuidados sino por la gloria y las recompensas que producen, y de los que debe huirse porque cuestan demasiado a la naturaleza» (*ibíd.*) Los comportamientos prosociales, que él llama comportamientos justos y virtuosos, son escasos en la interacción social, y existen porque se crearon las leyes formales, las normas sociales y la autoridad política, con el objeto de hacer posibles unos estándares de cooperación suficientes para que la sociedad sea viable²⁵. La cooperación con los

²⁴ El enfoque que Platón expone sobre el egoísmo resulta representativo del parecer de otras corrientes filosóficas de orientación individualista, como las que parten de Demócrito y Epicuro.

²⁵ Como advierte Popper (1991[1979]: 107),

... para Platón -y para la mayoría de los platónicos- no es posible la existencia de un individualismo altruista. Según Platón, la única alternativa fuera del colectivismo es el egoísmo, pues simplemente

demás, en términos platónicos, es un comportamiento interesado, pues lo que verdaderamente motiva al ser humano (a lo que le impelen sus pasiones) es el deseo de acumular recursos y poder para sí mismo, sin tener en cuenta la repercusión de ese deseo en los demás. Según Platón no puede ser de otro modo:

Demos al hombre de bien y al hombre malo un poder igual para hacer todo lo que quieran; sigámoslos, y veamos a dónde conduce la pasión al uno y al otro. No tardaremos en sorprender al hombre de bien siguiendo los pasos del hombre malo, arrastrado como él por el deseo de adquirir sin cesar más y más, deseo a cuyo cumplimiento aspira toda la naturaleza como a una cosa buena en sí, pero que la ley reprime y limita por la fuerza. (*La República*, II, 2, 2; 359b-359d)²⁶.

Aristóteles y las escuelas post-aristotélicas, destacadamente el estoicismo, respondieron a esta concepción de la naturaleza humana y de las instituciones sociales. Los conceptos clave son los de vida según la naturaleza, virtud y amistad; y, dado que el sentimiento de identificación que lleva a cooperar en la amistad no se extiende a los extraños y desconocidos, entonces, la conciencia de pertenecer a un linaje cultural común es el factor natural que posibilita cooperar (aunque sea sólo a modo de tolerancia y no-agresión) con los que no pertenecen al círculo de vinculaciones más fuertes.

Explica Montaigne en los *Ensayos*, que en el sistema de pensamiento aristotélico las relaciones sociales y la amistad ocupan un lugar tan prominente en la definición de la naturaleza humana que la amistad, considerada una virtud, se plantea incluso como un asunto de orden político, dadas las implicaciones de las conductas y sentimientos amistosos en la estabilidad y el orden social. Además, en último extremo, aunque no menos importante para el aristotelismo, la virtud de la amistad determina la felicidad concebida como un bien común:

Parece que nada hay a que la naturaleza nos haya encaminado tanto como al trato social. Aristóteles asegura que los buenos legisladores han cuidado más de la amistad que de la

identifica todo tipo de altruismo con el colectivismo y cualquier tipo de individualismo con el egoísmo.

Téngase en cuenta que Popper considera colectivista el modelo de Estado propuesto por Platón. Así pues, para Platón no resulta posible que los individuos actúen altruistamente de manera espontánea, dado que son egoístas por naturaleza.

²⁶ Evolutivamente, la estrategia “hombre de mal” tiene una eficacia superior que la “hombre de bien”, por lo que la segunda está condenada a la extinción. Si la diferencia de las eficacias es muy elevada, el egoísmo y la no-cooperación extinguen rápidamente a la cooperación. Es como el encuentro de un halcón con una paloma: si la paloma no tiene habilidades suficientes para escapar, el halcón la matará. El problema es que una sociedad poblada únicamente por halcones no es sostenible y se dirige hacia su desaparición. Véase Maynard Smith y Price (1973).

justicia. El último extremo de la perfección en las relaciones que ligan a los humanos reside en la amistad; por lo general, todas las simpatías que el amor, el interés y la necesidad privada o pública forjan y sostienen, son tanto menos generosas, tanto menos amistades, cuanto que a ellas se unen otros fines distintos a los de la amistad, considerada en sí misma. (Montaigne, *Ensayos*, I, xxvii, “De la amistad”).

En la *Política* (1253a) Aristóteles llega a afirmar que el hombre es un animal más social aún que las abejas y otros animales gregarios. La vida social humana depende de las virtudes, y su concepto de vida buena es la vida según la virtud, que consiste en lo contrario al inmoderado egoísmo²⁷. Este último es una forma de vida patológica y autodestructiva según el filósofo. La virtud, entendida en el sentido indicado, es una conducta natural en el ser humano (puesto que es animal social y racional antes que político), que permite cooperar con parientes biológicos y afines culturales, evitando conflictos nocivos para el propio individuo. Una de las virtudes más importantes para Aristóteles es la amistad (*philia*), que él presenta como una poderosa disposición que incita a cooperar y a formar los grupos naturales previos a la organización política de la sociedad. La red de vinculación fuerte, que se estructura en torno a la familia nuclear y los parientes, constituye la base social fundamental de la cooperación en el aristotelismo. La amistad se halla de forma natural en la familia, que es la simiente de la comunidad y de los comportamientos y sentimientos que sirven de base a la vida política²⁸. Los *amigos* para Aristóteles son un amplio conjunto de relaciones vinculatorias que incluyen lo que hoy día llamaríamos conocidos, vecinos, compañeros de tareas, los clientes y proveedores habituales o los socios en las empresas, y además y principalmente, aquellas personas a las que se ama, como los hijos, la pareja y los parientes, así como los amigos íntimos²⁹. ¿Qué es lo que caracteriza a este tipo de relación amistosa? Queda claro que con los amigos no existe el egoísmo, y la propia relación no está sujeta a intereses instrumentales de tipo utilitario:

²⁷ Gintis y Khurana (2008: 316) definen las virtudes de carácter como: «regularidades de comportamiento éticamente deseables que los individuos valoran por sí mismas, pero que a la vez tienen la propiedad de facilitar la cooperación y propiciar resultados socialmente eficientes.»

²⁸ En la *Ética Eudemia* (1242a) afirma:

Pues el hombre no es sólo un animal político, sino también familiar, y no se empareja como los demás (animales), alguna vez y con un hombre o mujer cualquiera, sino que el hombre es, propiamente, no un animal solitario, sino *comunitario con los que son sus parientes por naturaleza*; así pues, habrá una comunidad y alguna forma de lo justo aunque no haya polis, pues la familia es una forma de amistad. (Cursivas y paréntesis añadidos).

²⁹ Ben-Porath (1980) denomina *F-connection* (*families, friends, firms*) a ese conglomerado de relaciones de alta confianza y fuerte vinculación.

... consideramos amar el compartir sufrimiento con el que sufre, no por alguna otra razón, como los esclavos respecto de sus amos, (que comparten su sufrimiento) porque cuando éstos sufren son duros, sino por razón de ellos mismos, como (hacen) las madres con sus hijos y las aves que comparten sus dolores. Pues el amigo quiere, ante todo, no sólo padecer dolor junto al amigo, sino también el mismo dolor (...) Y el mismo razonamiento (cabe) respecto del alegrarse, pues es propio del amigo alegrarse no por alguna otra razón, sino porque el otro se alegra. (*Ética Eudemia*: 1240b; paréntesis añadidos).

En la amistad aristotélica la identificación entre las personas llega a ser tan motivadora, que la esfera cognitiva de lo concerniente a cada cual se extiende hacia una conciencia común de asuntos incumbentes. A esta conciencia compartida la podemos llamar espíritu comunitario o interés por la comunidad. En palabras de Montaigne,

En la amistad de que yo hablo, las almas se enlazan y confunden una con otra por modo tan íntimo que se borra y no hay medio de reconocer la trama que las une. (Montaigne, *Ensayos*, I, xxvii, “De la amistad”).

Aunque puede resultar chocante en la mentalidad de nuestros días, Aristóteles incluía dentro de las relaciones virtuosas de amistad identificatoria a las mujeres e incluso a los esclavos. En la cultura de su época, y de épocas ampliamente posteriores, se creyó en la inferioridad natural de la mujer respecto al hombre, así como en la diferente distribución de la fuerza y del talento entre los hombres. Esto justificaba relaciones simbióticas de subordinación, dominio y propiedad, que sin embargo no son incompatibles, según el Estagirita, con los afectos vinculatorios y la identificación con el bienestar y los intereses de mujeres y esclavos, motivaciones que son definitorias de la amistad libre entre personas virtuosas:

El esclavo es una parte del amo, como si fuera una parte animada, y separada, de su cuerpo. Por eso entre el esclavo y el señor, que por naturaleza son dignos de su condición, existe un cierto interés común y una amistad recíproca. En cambio, entre los que no se da tal relación, sino que lo son por convención y forzados, sucede lo contrario. (*Política*: 1225b).

¿Y qué características tienen los amigos para que surja la amistad? El origen de estas vinculaciones naturales, que se manifiestan en una esfera cognitiva compartida de las experiencias personales, es la semejanza:

... el amigo quiere ser, como dice el proverbio (...) otro yo. Pero está separado (del amigo) y es difícil llegar a ser uno; sin embargo, es lo más afín desde el punto de vista de la naturaleza: así, uno es semejante (al amigo) en el cuerpo, otro en el alma, otro en una parte de éstos y otro (en otra). Pero, no obstante, el amigo quiere ser como un yo separado. Por

tanto, percibir al amigo es necesariamente, de alguna manera, percibirse a uno mismo, y conocer al amigo, conocerse de alguna manera a uno mismo. (*Ética Eudemia*: 1245a; paréntesis añadidos).

Pero, ¿cuán semejantes han de percibirse dos individuos para sentirse vinculados por esta genuina amistad que desemboca en conductas identificatorias, es decir, en un interés espontáneo por cooperar? Aristóteles explica que entre desconocidos que no se reconocen como semejantes la cooperación no aflora naturalmente. En estas relaciones con no-afines, «como (las que se producen entre) los que venden y los que compran», puede darse una amistad de orden inferior, que es la amistad motivada por la utilidad, o sea, por la conveniencia recíproca. La cooperación instrumental es la que sostiene la sociedad política. Tanto en el comercio como en la convivencia política la cooperación no es fruto del amor ni del sentido de pertenencia (ingredientes de la semejanza que fragua las identidades grupales más fuertes), sino del acuerdo y del contrato. Tales amistades son utilitarias, y por lo tanto impulsadas por el egoísmo. En las relaciones motivadas por el egoísmo surgen disputas y recriminaciones, debido a que la confianza con extraños «es contraria a la naturaleza» (*Ética Eudemia*: 1243a). Estamos pues en una situación parecida a la descrita por Platón, y de ahí la necesidad del contrato privado, las leyes públicas y la administración de justicia, es decir, el refuerzo de la cooperación por terceros.

En la visión del mundo de las culturas influidas por el estoicismo y el cristianismo la base de la cooperación con todos los desconocidos, independientemente de la raza, etnia, o pertenencia a un orden político formal, requería algo más que la coerción de la autoridad política en pos de las leyes y del cumplimiento de los acuerdos adoptados. Evidentemente, antes de eso han de producirse los acuerdos, y antes de los acuerdos una aproximación confiable. En efecto, existe un ámbito de aproximación, previo al acuerdo entre egoístas, para el cual es necesario que los individuos se toleren y respeten de acuerdo a una cooperación “de mínimos”. Sin esa tolerancia cooperativa básica es imposible participar en interacciones instrumentales posteriores, dado que sin ella surgiría la desconfianza incluso hacia los magistrados que administran las leyes, y hacia el propio sentido de la ley.

¿De dónde procede la motivación que induce a esa tolerancia? El pensamiento estoico, antes que el cristianismo paulino y agustiniano, puso el acento en la idea de hermandad universal de todos los humanos. Es necesario que las personas *se reconozcan* como semejantes en una identidad humana, para que no se depreden ni se exploten unos a otros. Un humano, precisamente porque es humano, no puede dañar o destruir sin más a otro individuo que es como él, una copia más o menos semejante de sí mismo. Si lo hiciera, su conducta estaría denotando una “falta de humanidad” que señala dudas sobre el propio carácter humano de sus instintos. El mero reconocimiento de una identidad humana ha de activar las emociones, propias del ser humano, suficientes para tolerar la presencia del otro aunque no existan amor ni otros sentimientos de identificación filial o cultural. Un individuo que no es capaz de aceptar la presencia del otro sin cumplir unas normas de conducta básicas de humanidad no merece ser considerado humano, del mismo modo que unas tradiciones culturales agresivas y violentas, contrarias a la interacción pacífica y al orden cooperativo, son barbarie impropia de la “humanidad” del hombre. Por eso la ética estoica hace tanto hincapié en la dignidad intrínseca del ser humano, un animal, para los estoicos, que no sólo vive en sociedad por instinto, sino por el dictado de la razón, y la razón, que es la naturaleza esencial compartida por toda la humanidad, invita a la relación social, a la benevolencia y la fraternidad con todos los hombres. Como explica Copleston (2011: I-343) en su análisis de la ética estoica:

La división de la humanidad en naciones hostiles entre sí es un absurdo: el sabio es ciudadano, no de este o de aquel Estado particular, sino del mundo. Partiendo de esta base, síguese que todos los hombres tienen derecho a nuestra benevolencia, que los esclavos tienen también sus derechos y hasta los enemigos tienen derecho a nuestro perdón y a nuestra compasión.

El giro dado por el estoicismo, y después por el cristianismo, supone conferir al ser humano una dignidad moral especial respecto a los demás seres vivos, y esta dignidad se asocia a un modelo de comportamiento respetuoso y cooperativo con aquellos que se comportan tal y como se supone ha de hacerlo un individuo humano, quien por naturaleza lleva en sí la conciencia de quién es. En este sentido Edgar Morin (2004: 79 y 85) presenta la hominización como “humanización”, y para él la humanización tiene un significado psicológico de alteridad. La alteridad, un rasgo genuinamente humano, capacita al hombre para «consagrar su Yo a un Nosotros y a un

Tú». Tal capacidad, sin la cual es imposible el altruismo moral que caracteriza a la especie humana, procede de que al podernos reconocer como sujetos, así también somos capaces de reconocer la subjetividad de los que son como nosotros, y en eso consiste lo humano de la hominización. Lo inhumano, por el contrario, surge cuando el sujeto deja de ver reflejada en el otro su propia humanidad.

Estos argumentos, típicamente estoicos y cristianos, son la base del derecho natural en los juristas y teólogos europeos que crearon el derecho de gentes, precursor del moderno derecho internacional y de los tratados internacionales en materia de derechos humanos (Tierney 1997). El acervo intelectual de estos avances institucionales es una innovación cultural que persiste en nuestros días. La cuestión aquí relevante es reconocer las señales que permiten admitir a un animal dentro de la categoría de lo humano, y por lo tanto perteneciente a la gran hermandad conespecífica de los hombres³⁰. El origen, sin embargo, vuelve a ser aristotélico: el Estagirita caracteriza a los humanos como razonadores prácticos, capaces de practicar la *phronēsis* y por ello inteligentes, pero dado que reconocía señales de inteligencia también en otros animales, por ello, la capacidad de prever acontecimientos y anticiparse de manera estratégica, planificar, etc., no podían ser rasgos exclusivos del animal hombre³¹. Esto le condujo a designar como atributo específicamente humano el de la virtud, o sea, la capacidad de actuar de acuerdo al bien común según los dictados de la conciencia moral, una cualidad racional específica, según él, de la psicología humana. La cooperación, y no la destrucción mutua propia de los enemigos, es una característica de los individuos que practican la virtud, y como es sabido, en el eudemonismo aristotélico sólo es posible la felicidad llevando a cabo una vida virtuosa, es decir, lo que él llama una vida buena (Foot 2002: 173-175). Así lo explica Alasdair MacIntyre en *Tras la virtud*:

Lo que me enseña la educación en las virtudes es que mi bien como hombre es el mismo que el bien de aquellos otros que constituyen conmigo la comunidad humana. No puedo

³⁰ Explica Stout (2008: 162) que la cuestión crucial a la hora de enjuiciar el trato dado a otro guarda relación con el hecho de si ese otro es considerado «un miembro de la comunidad relevante cuyos intereses son objeto de consideración». Téngase en cuenta que, como ya defendiera Darwin (1972[1871]: I, iv), la consideración de pertenencia a la comunidad es una base firme del instinto de simpatía, y sobre este instinto se erige la moralidad de las comunidades humanas.

³¹ La inteligencia animal ha sido reconocida desde tiempos muy pretéritos, por ejemplo por Aristóteles y santo Tomás de Aquino. Darwin (1972[1871]: I, iii. p. 64) resalta la inteligencia animal como un rasgo definitorio de todas las especies, si bien dentro de una muy amplia graduación.

perseguir mi bien de ninguna manera que necesariamente sea antagónica del tuyo, porque *el* bien no es ni peculiarmente mío ni tuyo, ni lo bueno es propiedad privada. De aquí la definición aristotélica de amistad, la forma fundamental de relación humana, en términos de bienes que se comparten. Así, en el mundo antiguo y el medieval, el egoísta es alguien que ha cometido un error fundamental acerca de dónde reside su propio bien y por ello se autoexcluye de las relaciones humanas. (MacIntyre 2004[1984]: 281).

Los estoicos, frente a Aristóteles, prefirieron definir al ser humano no por su racionalidad, sino por ser un animal cuyo destino ineludible es la vida en comunidad. Pero ya que otros animales también viven agrupados, el rasgo distintivo adicional de la especie humana fue para el estoicismo un atributo metafísico: el *Logos* (testigo de un alma inmortal) habita en todos los hombres, y por ello disponer de conciencia y conocimiento -es decir, de una cosmovisión intersubjetiva- es la señal de que se es humano³². Una mente capaz de controlar los instintos agresivos, la conciencia moral en la relación con sus semejantes, empatía, cooperación y vida en comunidad, son características que fueron atribuidas a la especie humana por la escuela estoica³³. Todos estos atributos que igualan a los hombres tienen su raíz en el parentesco que comparten, un parentesco arraigado en su pertenencia al linaje divino:

Dado el parentesco entre la divinidad y los hombres, ¿qué otra cosa les queda a éstos, sino lo que decía Sócrates: al que pregunta “¿De dónde eres?”, no responderle nunca “ateniense” o “corintio”, sino “ciudadano del mundo” (...) ¿Por qué no va a decir que es ciudadano del mundo?, ¿por qué no que es hijo de la divinidad? (Epicteto, *Disertaciones*: I, IX).

Cleantes, quien sucedió a Zenón tras la muerte de éste en la dirección de la escuela estoica, escribió un himno a Zeus que, según precisa Kenny (2005: 132) era conocido por san Pablo, pues se sabe que lo citó en su predicación en Atenas:

¡Oh Dios gloriosísimo, que tantos nombres tienes,
Gran Rey de la Naturaleza, idéntico a lo largo de años sin fin;
Omnipotencia, que con tu justo decreto
lo riges todo! ¡Salve, o Zeus, según a Ti
es digno que por doquier tus creaturas te aclamen!
Hijos tuyos somos, los únicos que, entre todos los seres
que sobre la anchosura tierra se agitan, llevamos
a todas partes, en nosotros, tu imagen. (Citado en Copleston 2011: I-337).

³² Dice Marco Aurelio en las *Enseñanzas para una conducta moral*: «Entra en el interior de los demás y permíteles entrar dentro de ti» (VIII. 61).

³³ Véase Copleston (2011: I, cap. 36).

En efecto, este énfasis puesto en la hermandad universal de los hombres lo encontramos igualmente en san Pablo cuando se refiere a la caridad en la *Primera epístola a los corintios* (1 Co 13, 1-13), y es uno de los puntales del pensamiento cristiano posterior. Es el caso, por ejemplo, de la doctrina agustiniana sobre el amor caritativo³⁴. Así, el único tipo de amor no-místico, sino mundano, que según san Agustín permite tratar al prójimo como prójimo, sin las arbitrariedades propias de otros amores posesivos (concupiscentes, y por lo tanto egoístas y conflictivos), y sin el desapasionamiento místico respecto al mundo (un mundo que ineludiblemente incluye al hombre), es el que surge de los sentimientos asociados a la conciencia de que todos los hombres son hermanos, al menos desde la perspectiva de que todos son descendientes de un mismo linaje. De acuerdo a san Agustín los humanos somos descendientes de Adán, por lo tanto todos (cristianos o no) nacidos pecadores y mortales. Lo que nos une en cuanto a humanos y nos invita a cooperar no es sólo una misma identidad en la fe ni una semejanza étnico-cultural, sino el sentimiento de parentesco, un afecto vinculador que surge de conocer la semejanza esencial: la de compartir el mismo origen y la misma naturaleza, que avocan a un mismo destino que es la degradación y la muerte. San Pablo, en su *Carta a los Efesios*, acude igualmente a una metáfora que se apoya en las ideas de unidad espiritual y linaje común:

... ya no sois extranjeros o advenedizos. Sois conciudadanos de un pueblo consagrado, sois familia de Dios, sois piedras de un edificio construido sobre el cimiento de los apóstoles y los profetas. (Ef 2, 19-20).

En los propios *Evangelios* se cita un pasaje en el que Jesús insta a sus seguidores a una compasión universal, más allá de las antepuestas simpatías y apegos naturales por los que las gentes se vinculan a su círculo estrecho de familiares y amigos. El amor a Dios requiere amarlo en el desconocido que sufre necesidad. En eso consiste la fraternidad y el amor-caridad que define a los justos que recibirán como premio la vida eterna:

Porque estuve hambriento, y vosotros me disteis de comer; estuve sediento, y me disteis de beber; llegué como un extraño, y me recibisteis en vuestra casa; no tenía ropa y me la disteis; estuve enfermo, y me visitasteis; en la cárcel fuisteis a verme (...) Os aseguro que todo lo que hayáis hecho en favor del más pequeño de mis hermanos, a mí me lo habéis hecho. (Mt 25, 35-40).

³⁴ La explicación que aquí se da está basada en el texto de Hannah Arendt (2001[1929]).

Se trata de otra forma de mencionar el sentido de pertenencia, la identificación con la familia cultural, y la inevitable interdependencia caritativa, repleta de compasión, que generan las penalidades y tribulaciones, la enfermedad, la vejez y la muerte. Estas son las mismas bases del amor espontáneo que defendió Aristóteles, solo que en el caso de san Agustín (y de san Pablo), al igual que los estoicos, la caridad abarca a todo el género humano y no sólo a las personas más próximas y cotidianas.

Hemos traído a colación a los pensadores cristianos porque de acuerdo a Eccles (1996: 115) el nacimiento de las grandes religiones trajo consigo la institucionalización de la enseñanza del comportamiento altruista, y ello a pesar de la paradójica agresividad mostrada históricamente por las iglesias en nombre de la religión. No obstante, una innovación cultural que contribuye a extender la cooperación fuera de las micro-redes de vinculación fuerte, y a considerar afines también a aquéllos que antes eran considerados salvajes o enemigos, no sólo contribuye a extender la cooperación, sino que también mejora la calidad de la misma. Obviamente, esta innovación cultural aporta ventajas evolutivas a quienes la asimilan³⁵. Esta fue, entre otras, una causa del éxito de la propagación del cristianismo frente a otras religiones de la Roma imperial:

... aún más revolucionario fue el principio de que el amor y la caridad del cristianismo se deben extender más allá de las fronteras familiares y tribales (...) En realidad, el amor y la caridad se deben extender incluso más allá de la comunidad cristiana (...) y, en efecto, fue la base cultural para la revitalización de un mundo romano que se hallaba gemebundo entre una multitud de miserias. (Stark 2001: 231).

Las tradiciones culturales y las ideas acerca del carácter definitorio de lo humano se discutieron desde la Baja Edad Media hasta el Renacimiento, pero en realidad el momento crucial es cuando los europeos en su emigración hacia Occidente por el Atlántico se encontraron con humanos que hacía 15.000-20.000 años se había perdido contacto. Es lo que Peter Watson (2012) ha denominado *The great divide*. Después de tanto tiempo, ¿se les podía considerar humanos? La polémica en aquel momento se planteó en términos de explicar a qué Tribu de Israel pertenecían, si es que

³⁵ Para Mises (2009[1949]: 199-202) los sentimientos de solidaridad y atracción cooperativa que se dan en los grupos humanos (desde la familia hasta la nación, e incluso la completa humanidad) nada tienen que ver con la consanguinidad, sino con el apego resultante a las estructuras sociales cooperativas que son benéficas para la supervivencia y reproducción de los individuos. Estas estructuras no surgen espontáneamente, sino que resultan de construcciones ideológico-culturales que se reproducen mediante creencias.

pertenecían a alguna. El hecho de que los clérigos españoles, con Bartolomé de las Casas como proponente más destacado, los identificasen como nietos de Noé y descendientes de alguna de las Tribus de Israel les otorgaba la naturaleza de humanos.

La polémica de la humanidad de los indígenas del Nuevo Mundo es conocida como la *polémica de los naturales*³⁶. Fue la sostenida por Bartolomé de las Casas y Juan Ginés de Sepúlveda en los años 1550-1551. Según Elliott (1996[1963]: 75-86), lo que estaba en liza no era decretar sobre si los nativos eran hombres, o por el contrario animales a domesticar y esclavizar³⁷.



Figura 1.1.- Esta ilustración que aparece en la edición de 1874 de la *Antropogenie* de Ernst Haeckel contiene una concepción racista de la evolución, en la que se representa al “salvaje” como un animal más semejante a otros simios antropoides que a un civilizado *Homo sapiens* occidental. Fuente: Gould (2010[1977]: 241).

³⁶ Para un estudio histórico véase Abulafia (2009).

³⁷ Véanse figuras 1.1., 1.2. y 1.3., en las que se muestra gráficamente el trasfondo de la polémica a lo largo del tiempo.

La naturaleza humana de los indios había quedado reconocida por la bula papal *Sublimis Deus* de 1537, en la que Paulo III declaraba el derecho de los indios a la libertad y la propiedad, así como a no imponerles el cristianismo de manera violenta. En realidad el debate de la denominada Comisión de Valladolid giró en torno a si los indígenas son de naturaleza inferior o no a los conquistadores, pues de ser inferiores estaba justificada la guerra de conquista y el sometimiento a la autoridad tutelar de los españoles según la doctrina aristotélica de la esclavitud natural (este fue el principal argumento de Ginés de Sepúlveda).

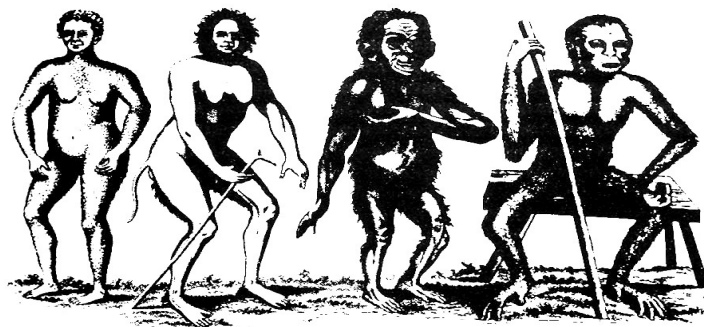


Figura 1.2.- Criaturas antropoides del imaginario sobre el salvajismo. Dibujo de un alumno de Linné, 1760. Fuente: Formigari (1984: 49).

Años antes, en 1511, ante el trato vejatorio propinado a los indios, el dominico Antonio Montesinos había exclamado públicamente en la isla La Española, y ante una indignada audiencia de colonos: «¿Acaso no son hombres estos indios? ¿No tienen almas racionales? ¿No sois obligados a amarlos como a vosotros mismos?» (Citado en Elliot 1996[1963]: 82).

De las Casas, seguidor de las ideas de Montesinos y de Francisco de Vitoria, reclamó el respeto a la dignidad de los indígenas, su derecho a no ser sometidos, tutelados ni subordinados, y a ser tratados en condiciones de igualdad. Según su punto de vista, la razón de los derechos de los indios estriba en que son tan humanos como los europeos en todas las manifestaciones admitidas en su época sobre la “humanidad” humana: portadores de una cultura, habitantes de ciudades, dotados de instituciones político-jurídicas, propiedad, comercio, etc., y en cuanto a sus actos de crueldad, no muy diferentes a los observados en el Viejo Mundo.



Figura 1.3.- Imagen estereotipada europea del salvaje en la Europa del s. XVI. Es representado como animal monstruoso no humano. Fuente: Formigari (1984: 89).

Los alegatos de Ginés de Sepúlveda, sin embargo, acabaron decantando su victoria en la polémica. Los más determinantes fueron aquellos relacionados con la crueldad manifestada en algunos comportamientos indígenas hacia sus semejantes, como es la realización de sacrificios humanos en ceremonias religiosas, o el canibalismo³⁸. Asimismo, consideró incivilizada la carencia de propiedad privada en algunas de estas sociedades. Estas conductas, impropias de un ser humano “civilizado”, sirvieron de justificación para el dominio por la fuerza, con propósito “tutelar” y reeducador, por parte de las potencias europeas en América, una actitud que fue acompañada de un espíritu de cruzada contra el mal que caracterizó tanto la inicial colonización católica del Nuevo Mundo, como la anglicana y puritana posteriores de Nueva Inglaterra (Cañizares-Esguerra 2008).

³⁸ La costumbre de la antropofagia ha sido tradicionalmente considerada por las culturas occidentales modernas un indicio de salvajismo pre-humano. En *La Enciclopedia* (1777) se dice que: «Una gran parte de América está poblada de salvajes, la mayor parte feroces, que se alimentan de carne humana.» (Citado en Iglesias 2006: 422-423).

1.2.2.- El egoísmo y el altruismo en los debates de la Ilustración previos al surgimiento de la economía política

Las dos concepciones históricas sobre la naturaleza humana, la prosocial y la egoísta, volvieron a resurgir en los debates post-renacentistas dando lugar a otras visiones enfrentadas sobre la naturaleza y el papel a jugar por las instituciones (fundamentalmente las políticas) en la organización de las relaciones humanas. De la primera, que asume la bondad natural de los hombres, surgirán teorías políticas (y posteriormente económicas) coherentes con órdenes sociales relativamente armónicos y tolerantes, en los que no es preciso el control del individuo por parte del poder político. De la segunda, que parte del pesimismo antropológico, surgirán teorías políticas que concluyen en la necesidad de sistemas políticos autoritarios, e instituciones que contribuyan a controlar el egoísmo predatorio y oportunista.

El egoísmo es una tendencia innata para los defensores de la hipótesis de una naturaleza humana egoísta, y sólo mediante el control social y político es posible la cooperación. La matriz básica de este argumento reaparece en las visiones del iusnaturalismo post-renacentista que cristalizaron en una tradición de pensamiento muy influyente en la posterior economía política³⁹. Así, encontramos estos supuestos conductistas de manera destacada en las obras de Hobbes, Locke, Mandeville, Hume, y otros filósofos que aportaron los fundamentos filosóficos de la teoría utilitarista del comportamiento, como Bentham (MacPherson 1973; Watson 1966). En esencia, su principal argumento es que los hombres tienden de manera natural a utilizarse interesadamente unos a otros, pues unos para otros carecen de valor intrínseco. Por lo tanto no existe genuina benevolencia ni otros afectos prosociales; sólo hay intereses egoístas en la interacción humana⁴⁰. Una arquitectura pasional y psicológica egoístas, y la disposición de la capacidad natural para el razonamiento práctico y estratégico, son los dos axiomas centrales de esta tradición filosófica que ha inspirado el pensamiento jurídico-económico, la filosofía moral y la ciencia política posteriores al s. XVII⁴¹. Tal y

³⁹ Véanse Padgen (2001: 21-24), y Tenenti (1999: 174 y 186).

⁴⁰ Como precisa Fukuyama (2000: 214) parafraseando a la jurista M. A. Glendon, la interacción de individuos sin inclinación natural por la sociedad viene a ser una agregación de iniciativas individuales conjuntas llevada a cabo como medio para lograr los respectivos objetivos individuales. Esta metáfora de individuos motivados exclusivamente de esta manera atomista es lo que Glendon denomina “poseedores solitarios de derechos”.

⁴¹ Véanse MacIntyre (2001a); Hausman (1998); Gauthier (1994[1986]); Macpherson (1973).

como escribiera Bertrand Russell (1982[1918]: 162), «el hombre es por naturaleza un competidor, un ser adquisitivo y más o menos belicoso». El instinto hace al hombre un ser competidor, y al ser además inteligente, elige reprimir su agresividad porque le conviene; por eso coopera con otros en la creación de reglas morales y jurídicas que, limitando su agresivo egoísmo posesivo, permitan la cooperación a la par que lo hacen minimizando la auto-represión de esos instintos agresivos y posesivos:

El argumento ético-económico requiere la aceptación inicial de una visión escéptica o pesimista de la naturaleza humana. Se reconoce que el egoísmo, ampliamente concebido, es una gran fuerza motivadora en toda actividad humana; se supone que la acción humana, si no está limitada por restricciones éticas o morales, está dirigida más naturalmente hacia la consecución del interés individual o privado. Esta visión de la naturaleza humana es, desde luego, esencialmente la adoptada por los filósofos utilitaristas. De esto se deduce directamente que el ser humano individual debe sufrir algún esfuerzo para restringir sus “pasiones” y que debe actuar de acuerdo con principios éticos o morales siempre que las circunstancias sociales y las costumbres dicten alguna desviación de los intereses privados. Tal esfuerzo, como todo esfuerzo, es escaso: es decir, es económico. En consecuencia, debería economizarse su empleo. Dentro de lo posible, se deberían desarrollar instituciones y limitaciones legales que ordenaran la búsqueda del beneficio individual de tal modo que lo hagan consistente con, en vez de contrario a, la obtención de los objetivos del grupo como todo. Sobre estos fundamentos psicológicos y éticos se basa la teoría de mercados o la organización competitiva de la actividad económica. (Buchanan y Tullock 1991[1962]: 327).

Como en el pensamiento de Platón, la cooperación se circunscribe a la conveniencia egoísta del individuo, y no cabe una cooperación sin previo cálculo neto de los intereses personales. Así, los intercambios libres entre los individuos han de ser necesariamente ventajosos *a priori* para quienes comercian. La economía moderna, de este modo, está concebida como una ciencia de la razón práctica, o elección racional, en la que los comportamientos son instrumentales al servicio del propio interés. Los grupos humanos, en los que por su propia naturaleza se coopera, son estructuras sociales pareto-óptimas respecto a otras alternativas de organización, pero que conllevan el coste de ser regulados y gobernados para evitar la tragedia de los comunes⁴².

⁴² La formulación más conocida de esta teoría en la actualidad es la del ecólogo humano Garret Hardin (1968), si bien existen numerosas formulaciones anteriores que se remontan al propio Aristóteles en la *Política*. La “tragedia de los bienes públicos o comunales” (tragedia de los comunes) consiste en que «el individuo no considera los costes externos cuando emprende el cálculo económico» (Bagus 2004: 127). Otra visión de la tragedia de los comunes consiste en presentarla como un juego de dilema del prisionero en el que participan más de dos jugadores (Ridley 1996: 230-231). El proceso que conduce a la tragedia se debe a que ante la indefinición, falta de refuerzo institucional o refuerzo defectuoso de los derechos de propiedad «no hay límites a la explotación de la propiedad no rigurosamente definida, y el incentivo es el de explotar esta propiedad cuanto antes para adelantarse a la explotación efectuada por los demás» (Bagus 2004: 137). En definitiva, se trata de un incentivo a la sobreexplotación de los recursos motivado por el afán de ganancia individual y la no internalización de externalidades.

¿Cómo se expresa la versión moderna de la controversia? La polémica tuvo su epicentro en las críticas de Joseph Butler a Hobbes y Locke, y se desarrolló en un doble sentido. Por un lado la polémica alude a la motivación de las acciones: Butler refutó el egoísmo racional propuesto por Hobbes y sus seguidores contra-argumentado con el ejemplo de su famosa “piedra”. Según Hobbes todo acto humano busca intencionalmente una satisfacción subjetiva, y por ello los seres humanos (que son individuos hedonistas compulsivos, además de racionales), son capaces de adaptarse estratégicamente a normas y estructuras sociales que tienen funcionalidad cooperativa, si bien no lo hacen por motivación intrínseca, sino con intención utilitaria. De esta manera, el egoísmo racional no es incompatible con la práctica de la cooperación social⁴³. Según Hobbes cooperamos por egoísmo, e incluso podemos ser fervientes filántropos, pero el altruismo en Hobbes siempre tiene una motivación egoísta, por ejemplo, calmar el desagradable sentimiento de la culpa, disfrutar la vanagloria, mejorar la reputación, un soterrado afán de dominio, o simplemente gozar de una forma particular de narcisismo filantrópico. Detrás de la filantropía hobbesiana siempre está una motivación hedonista, dado que para Hobbes el egoísmo utilitario es la única motivación humana (Watson 1966: 129)⁴⁴.

Butler niega estos supuestos, y responde contra-argumentando con la evidencia empírica de las bases instintivas (pre-rationales, e incluso inconscientes) de la conducta humana: no actuamos racionalmente buscando recompensas hedónicas o utilitarias, sino que actuamos impelidos por lo que hoy llamaríamos instintos o etograma, y el que obremos de este modo no es incompatible con un refuerzo hedónico como consecuencia no buscada de la acción. Concretamente Butler afirma que:

El hecho de que todos los apetitos y pasiones se dirigen hacia los *elementos externos mismos*, que se diferencian del *placer que se obtiene de ellos*, se pone de manifiesto con lo siguiente: no podría existir este placer de no existir antes una correspondencia entre el elemento y la pasión; no podríamos disfrutar o deleitarnos más con un elemento que con el otro, más por comer un alimento que por ingerir una piedra, si no hubiese una inclinación o apetito hacia un elemento más que hacia otro⁴⁵.

⁴³ Véase Hobbes (1993[1651]: I; viii; x, xi y xiii). Una versión actual de las teorías hobbesianas es la de Gauthier (1994[1986]), que respalda una visión contractualista de la moral.

⁴⁴ Para un ampliación analítica de esta perspectiva véanse asimismo Gooddenough (2008), Trincado (2004) y Fontaine (1997, y 2000).

⁴⁵ Butler, en *Fifteen sermons upon human nature* (1726), citado en Sober y Wilson (2000: 243; cursivas en el original).

Una refutación parecida es la que hace Adam Smith en su ensayo “De los sentidos externos”⁴⁶. Argumenta el autor que los movimientos espontáneos del bebé recién nacido demuestran que en su naturaleza existen instintos que le conducen a actuar según lo que necesita en cada momento. El bebé se mueve sin necesidad de decidir conscientemente. Las sensaciones agradables que pueda experimentar no consisten en el resultado de una búsqueda premeditada del placer o la utilidad (un bebé desconoce las consecuencias de sus actos), sino la afirmación ontológica del organismo vivo, una afirmación que no es deliberada, calculada, ni dirigida, pero que consiste en una manifestación singular de la vida⁴⁷.

Lo que están enfatizando Butler y Smith es que existen automatismos instintivos que activan apetitos previos a toda deliberación racional, de manera que la sensación de hambre o las ganas de moverse focalizan en sí mismas la atención hacia una gama de objetos y comportamientos y no otra. La refutación de Butler y de Smith puede hacerse utilizando un lenguaje actual, como hace Frank (1988: 197) cuando explica que lo que motiva para comer algo, que es la sensación de tener apetito, viene explicado por el estado metabólico del cuerpo, es decir, fuerzas bioquímicas que motivan para actuar y cuyo manejo el individuo desconoce (no se dispone de información acerca de los pagos bioquímicos o energéticos asociados a la acción de comer). No obstante, se atiende espontáneamente a esas fuerzas motivadoras, y así haciéndolo se obtiene un resultado satisfactorio. Se podría decir, así pues, que no es preciso optimizar conscientemente ninguna ecuación; ya lo hace el propio organismo sin cálculo deliberado, igual que lo hace en un saltamontes o una amapola.

La trascendencia de estos argumentos apunta a negar el poder de la razón sobre los instintos, los cuales gobiernan determinísticamente la vida. Una prueba de ello es que se desea vivir no porque sea racional hacerlo, sino porque los instintos impelen a vivir. Por mucho que racionalmente se tenga consciencia del deseo de supervivencia, sin embargo no por tal conocimiento y la motivación asociada a él se vive eternamente⁴⁸. El

⁴⁶ El ensayo pertenece a los *Ensayos filosóficos*. El pasaje comentado está en Smith (1998[1795]: 170).

⁴⁷ Wilson y Daly (1997: 256-257) indican que existen argumentos suficientes para afirmar que la conducta de los neonatos es una programación genética que en su mayor parte está configurada para “seducir” a sus madres a través de la emisión señales que denotan calidad y buena salud.

⁴⁸ Del mismo modo, un órgano lesionado o enfermo no puede ser sustituido en un trasplante a no ser que sea aceptado por el organismo. Al fin y al cabo, no decide la razón si es más útil que el órgano

corolario del argumento puede ser expresado en terminología actual: el egoísmo racional es una falacia; el dolor es una señal de alarma que informa de un problema de supervivencia, y sirve para sobrevivir, igual que las sensaciones de placer, bienestar o júbilo son refuerzos de las conductas adaptadas para la supervivencia, y no motivos-*a priori*-para-actuar.

Reenfocando la funcionalidad de las sensaciones hedónicas, hoy día sabemos que tanto el placer como el dolor son señales que informan sobre el estado homeostático del organismo, y no objetivos finalistas en sí mismos del comportamiento humano (Damasio 2005). Así, el placer, el bienestar o la autoestima constituyen “recompensas” *ex post* que refuerzan una conducta. Pero no se realiza sanamente tal conducta buscando *ex ante* la incierta recompensa, sino que se hace porque tal comportamiento está codificado en procesos orgánicos congénitos. El estímulo externo o la imaginación activan el deseo vía *apetito*, no descuento racional del placer por venir; y de este modo, cuando se come o se bosteza se experimenta placer y sensaciones agradables, pero no es esa expectativa lo que impulsa la acción de comer o bostezar, sino el deseo de hacerlo, y ese deseo se activa mediante la conciencia del apetito, o motivación. En este sentido, el emotivismo de Hume resultó contundente en su defensa filosófica de la ineludible servidumbre que rinde la razón a las pasiones, y por ello igualmente el libre albedrío a los procesos físicos de la naturaleza:

Preguntad a un hombre *por qué hace ejercicio*, y os responderá que *porque desea conservar su salud*. Si le preguntáis entonces *por qué desea la salud*, inmediatamente os contestará que *porque la enfermedad es dolorosa*. Si lleváis vuestras inquisiciones más allá y deseáis que os dé una razón de *por qué odia el dolor*, es imposible que jamás pueda daros ninguna. Se trata de un fin último, y no puede ser referido a ningún otro objeto (...) Tiene que haber algo que sea deseable por sí mismo, debido a su inmediata concordancia o acuerdo con el sentimiento y afecto humanos. (Hume 1993[1751]: §112, V; cursivas en el original).

Este argumento antiutilitarista de los apetitos instintivos, que pueden ser en buena medida adaptados a las circunstancias sociales y culturales del momento a través de la enculturación, constituyó un tema central del discurso antihobbesiano de la Escuela Escocesa de filosofía moral. Especialmente clara se aprecia la tesis en la obra moral y psicológica tanto de Adam Smith como de Adam Ferguson. Para este último,

trasplantado sea rechazado o aceptado, sino que son los innatismos (procesos orgánicos) los que “dictaminan”.

que lo expone en sintonía con la tesis butleriana, la sociabilidad no es una cuestión circunstancial de ventajas calculadas por el raciocinio, sino que se trata de una genuina pulsión instintiva (que produce apetitos) en la especie humana⁴⁹. Así, los humanos son sociables por naturaleza, y además es posible fortalecer esta sociabilidad natural mediante la educación, dado que la cultura es un producto (un epifenómeno) complementario de la propia naturaleza social humana. Dice Adam Ferguson en su *Essay on the history of civil society*:

... el hombre, como el resto de los animales, tiene ciertas propiedades instintivas que son anteriores a la percepción del placer y del dolor, y que le llevarán a realizar muchas funciones (...) que tienen relación con sus semejantes. (Ferguson 1974[1767]: 68).

Así, de este modo, aun cuando el individuo busque el bien de otros y ese bien de otros acabe repercutiendo en su propio bien, no necesariamente lo hace motivado por este último propósito.

En un segundo bloque argumental la polémica moderna sobre el egoísmo alude a la epistemología de la identidad personal. Locke (1980[1690]: II, xxvii, 9) define la persona como sigue:

... un ser pensante inteligente, que tiene razón y reflexión, y puede considerarse a sí mismo como él mismo, la misma cosa pensante, en diferentes tiempos y lugares; lo que hace sólo mediante esta conciencia que es inseparable de pensar y, me parece, esencial al mismo: siendo imposible para nadie percibir sin percibir que percibe.

Queda patente por lo tanto que Locke plantea una teoría psicológica de la individuación, basándose en una teoría de la conciencia según la cual, como en Descartes, ser-equivale-a-conocer, y del conocimiento de que se conoce surge la primera persona del singular por la vía de las experiencias auto-atribuidas. Las experiencias conocidas se auto-atribuyen como si se adhirieran en “algo” que se muestra pasivo y receptivo a la penetración de información aferente (una *tabula rasa* mental). En esa operación las experiencias forman un nexo cognoscitivo que, por el hecho de ser conocido y experimentado en primera persona del singular, propicia que la información adherida

⁴⁹ La psicología social contemporánea se refiere a esa fuerte pulsión prosocial con el término “necesidad de pertenecer (a un grupo)” (*need to belong*) (véase Baumeister y Leary 1995). Sugden (2005 y 2009) denomina *fellow-feeling* a los sentimientos morales agradables relacionados con la compañía y la amistad.

sea considerada una especie de patrimonio del yo, constitutivo de una identidad personal en el tiempo⁵⁰. Estas experiencias son subjetivas en la medida que se da una cognición privada (una especie de propiedad o patrimonio epistemológico unificado, corriente continua y coherente de experiencias, o “haz de percepciones”, que dirá Hume):

Locke afirma que la conciencia es la apropiación de cosas y actos que incumben al yo y que son imputables a ese (yo) (...) que implican una relación de reconocimiento y apropiación de hechos e ideas de las cuales el yo es responsable. (Bartra 2006: 12-13; paréntesis añadido).

Así pues, en la tesis de Locke la conciencia de lo que ocurre es la explicación de que existe un sujeto, y el recuerdo histórico-biográfico de ese acontecer en primera persona es su identidad personal. Butler mostró su insatisfacción con esta teoría, y argumentó concretamente que: «nos debería parecer autoevidente que la conciencia de la identidad personal presupone la identidad personal, y por tanto, no puede constituirla» (Butler 1975[1736]: 100). Lo que está indicando es que conocer y recordar experiencias no implica “propiedad” de las mismas, es decir, falta el sujeto; o en palabras de Martín Lozano (1995: 76):

Todo recuerdo, aparente o verídico, remite a una experiencia pasada. El que afirma recordar algo cree que una de sus experiencias pasadas mantiene una relación peculiar con su recuerdo; de no ser así, no diría que está recordando, sino, tal vez, que está sufriendo una alucinación, o que ha descubierto que es clarividente.

De este modo, la crítica de Butler a Locke se centra en la que a su juicio es una errónea concepción del sujeto, ya que para el obispo la mera conciencia de una experiencia en el tiempo no implica que esa experiencia pertenezca a “alguien” que la conoce, es decir, constituye un error identificar la historia de experiencias con la del sujeto⁵¹. Según Butler el sujeto no puede ser una narración en primera persona de tipo

⁵⁰ Véase Locke (1980[1690], especialmente II, xxvii). Algunos estudios especializados sobre el asunto son los de Cuypers (1993: 225), Martín Lozano (1995: 76), y Sanfélix (1994: 268).

⁵¹ La teoría lockeana ha inspirado las aproximaciones de la psicología y neuropsicología contemporáneas al estudio de la identidad personal (p.ej. Jetten, *et. al.* 2010, y Klein, *et. al.* 1996). Sin embargo, y aunque todavía no se conocen con exactitud las estructuras funcionales del cerebro que permiten la fenomenología del yo, hay estudios que concluyen, como defiende Butler, que el yo es algo más que conciencia y memoria. Así, los pacientes con graves amnesias debidas a daños y lesiones neurológicas, como la enfermedad de Alzheimer, conservan, aun en estados muy avanzados de la amnesia, un autoconocimiento funcional de los rasgos más significativos de su personalidad, gustos y aversiones (Cohen-Mansfield, *et. al.* 2000; Klein, *et. al.* 2002). Y eso que el paciente

identificativo-posesivo, sino una entidad diferenciada (un alma inmaterial) que reconoce qué experiencias le conciernen y le afectan y cuáles no, o sea, un proceso activo de discriminación-de-lo-que-acontece-en-el-mundo⁵². Parafraseando a dos neurobiólogos actuales, las experiencias subjetivas (los *qualia*, en la jerga filosófica) y el yo, son las dos caras de la misma moneda:

Es evidente que no existen *qualia* flotando libremente sin ser experimentados por nadie, y resulta difícil imaginar un yo totalmente privado de *qualia*. (Ramachandran y Blakeslee 1999: 308).

Si se prescinde del contenido metafísico en el argumento de Butler, puede entenderse que a lo que él apunta es a que la sensación de lo que le ocurre *al* sujeto, diferenciada de manera clara y distinta respecto a lo que acontece a otros sujetos, no es un asunto solamente cognitivo, sino que previa y principalmente las sensaciones propias son *organizadas* y *sentidas* como propias de manera instintiva, o afectivo-intuitiva si utilizamos un lenguaje más psicológico. Así, mi dolor de cabeza es “mío” no porque lo conozca en primera persona, sino porque ese dolor “*me*” *afecta*; el dolor me está informando de que algo ocurre en relación con un asunto de “*mi*” *incumbencia*: el estado vital de “*mi*” organismo. Sólo asumiendo ésto es posible explicar las emociones y motivaciones del hombre real, como el amor y el odio, el propio interés, el compromiso, el orgullo y la culpa, el apego a los objetos, la responsabilidad de las acciones, etc., cuestiones difícilmente explicables con el *yo* meramente cognitivo que teoriza Locke⁵³.

considera definitivo de sí mismo (“cómo es”) se corresponde bastante bien con la descripción de personas sanas que lo conocen íntimamente desde mucho tiempo antes de padecer la enfermedad.

⁵² En la crítica a las teorías neo-lockeanas y neo-humeanas del sujeto, Edith Stein se opone a los alegatos filosóficos que definen al yo como una corriente mental indiferenciada de información. Su refutación contiene la esencia del argumento de Butler, y se refiere a que para ser consciente del acontecer cambiante, es necesario que exista algo previo y estable que lo observe:

... cada vivencia es esencialmente vivencia de un yo, y fenomenológicamente tampoco es separable en absoluto de él (...) (el yo) no es otra cosa que el sujeto del vivenciar viviente en el vivenciar. Así entendido, se torna carente de sentido la cuestión de si una vivencia es “*mía*” o de otro. (Stein 2004[1917]: 46; paréntesis añadido).

⁵³ Hume (1992[1739-1740]: I, iv, 6) también se dio cuenta del problema de la identidad personal lockeana, y por eso no fundamenta el yo en el conocimiento y en la memoria, sino en las categorías mentales de semejanza, continuidad y causalidad de una «impresión muy íntima» de permanencia existencial que, sin embargo, consiste en un espejismo. Pero Hume no atisba qué función cumple esa ilusión (o punto ciego) en la naturaleza humana, dado que no disponía de una teoría de la evolución biológica en la que engarzar su teoría de la consciencia. Su reflexión sobre el yo es filosófica, no científica.

En realidad las teorías empiristas de la identidad personal de Locke (y su colofón en Hume) son explicaciones cognitivas del acontecer, pero desconectadas de un arraigo funcional en la vida físico-orgánica (Gregory 2004). O sea, las teorías solamente cognitivas de la identidad no se refieren a un *yo* “encarnado” en un mundo real que se enfrenta a problemas de supervivencia y a relaciones con otros *yoes*⁵⁴. A diferencia del sujeto cognitivo lockeano, un *yo* encarnado es una función adaptativa de un organismo, y ese *yo* no sólo conoce acontecimientos, sino que ese conocimiento produce afectación, que le emociona, le conmueve y motiva para actuar. De entre las experiencias conocidas, el *yo* encarnado no se identifica igualmente con todas, sino que lo hace de manera selectiva, atribuyéndoselas como algo que le incumbe y que le predispone (sí o no, y en mayor o menor medida) para alguna meta biológica. El *yo* biológico sería una creencia derivada de la unificación de las respuestas emocionales del organismo y una memoria emocional (que permite expresar juicios de valor), en lugar de una acumulación de datos neutros (Damasio 2004).

Debido a esto, resulta claramente dudosa la teoría lockeana de la responsabilidad ante los propios actos, así como inexplicable en la coordenadas lockeanas el fenómeno de las identidades supra-individuales, aunque ambos aspectos son necesarios para la intersubjetividad, y sin ellos no habría vida social (y prosocial) propiamente humana⁵⁵. Aquí también Adam Ferguson se enfrenta a esta concepción informativa de la identidad personal, en la que difícilmente cabe la identificación con las experiencias e intereses de otros *yoes*, dado que falta el anclaje de las emociones y sentimientos del ser humano de carne y hueso. Un espíritu desencarnado, como el alma cartesiana o el *yo* lockeano, no podría comprender ni atribuir experiencias, apetitos o deseos a ningún cuerpo, ni el supuesto cuerpo “propio” ni tampoco los cuerpos de otros *yoes*⁵⁶. Además, muestra su

⁵⁴ Cuando hablamos de *yo* “encarnado” nos referimos a la experiencia mental como metarrepresentación de procesos vitales conscientes e inconscientes, que se producen en un organismo programado para sobrevivir y reproducirse, mediante patrones de acción fijos y patrones flexibles (Damasio 2005: 85).

⁵⁵ El fenómeno de la identidad supra-individual consistente en conocer las experiencias de otros *yoes*, e identificarse con ellas sin perder la referencia subjetiva de la individualidad. Hoy día han surgido diversos circunloquios para denominar esta operación psicológica radicada en la capacidad humana de empatizar: racionalidad colectiva y razonamiento de equipo (Colman, *et. al.* 2008a), “nostrismo”, intencionalidad colectiva y *we-rationality* (Schmid 2005 y 2007), pensamiento y preferencias de equipo (Sugden 1993 y 2000).

⁵⁶ Edith Stein explica la indisociable unidad de la mente (alma) al cuerpo vivo, y desde ahí a otras mentes vinculadas a otros cuerpos vivos, indicando que:

Gracias a esta fundamentación del alma en el cuerpo vivo, la empatía en individuos psicofísicos es posible sólo para un sujeto del mismo tipo. Un *yo* puro, por ejemplo, para el que no se constituye originariamente

desacuerdo con esas otras concepciones hedonistas del sujeto, como la de la tradición que conecta a Gassendi, los libertinos franceses, Hobbes y Mandeville. En efecto, para Ferguson (y para Adam Smith) no es correcto identificar, como hacen los hedonistas, el instinto de autoconservación con el amor a sí mismo (*amor sui*) y el propio interés. La expresión *amor sui*, que puede conducir, como propugna Hobbes, incluso a un excelso altruismo instrumental, viene a ser según los escoceses un error gramatical. Esto es así porque para ellos el amor es una emoción que por su propia naturaleza requiere reconocer la existencia subjetiva (y no sólo objetual) de un *tú* de carne y hueso además de la de un yo, y por esta razón los intereses de ese *tú* no pueden ser atendidos (solo y siempre) por motivaciones egoístas (como las del cálculo utilitario *a priori*). Cuando se ama, la atención puesta en los otros y sus intereses, la «compañía de sentimientos» que diría Smith, es un «principio de la naturaleza humana», o sea, un instinto (Smith 1997[1759]: I, iii, 1. p. 113). Por eso la experiencia de tal “compañía” es necesariamente previa al cálculo utilitario apriorístico del yo⁵⁷:

El amor es un sentimiento que lleva la atención de la mente más allá de sí misma, y su objeto se da en sentido de relación con algún semejante. (Ferguson 1974[1767]: 16).

1.2.3.- La polémica en la actualidad: la sociobiología, los neohobbesianos y los neobutlerianos

La polémica filosófica sobre el egoísmo ha llegado hasta nuestros días y se mantiene entre las tradiciones racionalistas-idealistas y el monismo sensualista-materialista. En esencia, la primera tradición niega el determinismo estricto de la acción humana en términos morales, enfatiza el poder de la razón y la voluntad, y se inclina por el monismo motivacional: la razón. La segunda tradición defiende la existencia de procesos deterministas en psicología y en la acción, y se inclina por una tipología plural de motivaciones para la acción: las que directamente proceden del sistema emotivo del organismo, y aquéllas otras en las que interviene la deliberación racional. Dependiendo de la consideración predominantemente egoísta o prosocial de las motivaciones para actuar (ya sean procesos deliberativos u emotivos), los discursos resultarán defensas racionalizadoras del egoísmo o del altruísmo. La controversia, no obstante, ha ido

un cuerpo vivo propio y una relación psicofísica, quizá podría tener dados objetos varios, pero no podría percibir cuerpos vivos, individuos vivientes. (Stein 2004[1917]: 106).

⁵⁷ Cuando decimos “previa” queremos referirnos a que es no-elegida, sino sentida y conocida por un impulso espontáneo.

desplazándose desde el ámbito de la interpretación filosófica hacia el de la contrastación experimental, y en este giro se encuentra el debate en la actualidad⁵⁸.

Un protagonismo destacado en este aspecto es el que ha jugado en la segunda mitad del s. XX la sociobiología, para la que, parafraseando a Edward O. Wilson, el altruismo consiste en el «problema teórico central» (Wilson 1980[1975]: 3). Es importante destacar que la discusión actual no se centra en la dicotomía maximalista egoísmo puro *versus* altruismo puro. Los teóricos científicos del altruismo no afirman que los seres humanos (ni los demás animales) sean altruistas en un sentido ingenuo⁵⁹. El biólogo George Williams (1996[1966]), por ejemplo, en su defensa del enfoque de la selección individual, sostiene que el altruismo, entendido como lo entiende la biología evolucionista moderna (comportamientos benéficos para otros organismos -emparentados o no- que conllevan un costo, en términos de éxito reproductivo, para el individuo que los realiza), no puede ser un comportamiento indiscriminado, ya que de ser así, los individuos egoístas tendrían más probabilidad que los altruistas de propagar sus genes a las siguientes generaciones y, en el devenir evolutivo de un número de generaciones, dependiendo de la intensidad del altruismo practicado y de la capacidad diferencial de reproducción de ambos rasgos, la población de los altruistas tendería a la extinción, surgiendo un ecosistema poblado sólo por egoístas⁶⁰.

En este sentido, López y Valdaliso (1999: 35-36), con una perspectiva conductual no sólo genotípica, sino también fenotípico-cultural, manifiestan que:

El altruismo más fuerte es el que implica sacrificar la vida o la reproducción en beneficio de otros, o de un organismo diferente u organización de tu especie, sin recibir nada a cambio ni en vida ni para tus descendientes. Este tipo de altruismo no tiene razón alguna para existir, tan sólo es una confusión.

⁵⁸ Una exposición filosófica de la polémica, decantada por la visión racionalista, se encuentra en Searle (2000). Dos aportaciones de la neurociencia al debate filosófico, que explican el funcionamiento neurológico de los procesos emotivos y deliberativos de decisión, son las de Cushman y Greene (2012), y Greene, *et. al.* (2004).

⁵⁹ La generosidad indiscriminada, sin sentido de medida ni dirección, ha sido calificada por Garrett Hardin (1977) de “altruismo promiscuo”. Este tipo de altruismo no existe en la naturaleza, aunque algunas religiones e ideologías lo prediquen como regla de oro.

⁶⁰ Esta explicación ha sido formalizada ampliamente por Maynard Smith (1982) y Parker (1984), y se refiere al concepto de estrategia evolutivamente estable: aquella que, una vez instalada en una población, no puede ser invadida por ninguna estrategia mutante.

Para ambos autores el altruismo que puede tener éxito evolutivo es en realidad una forma de egoísmo aplazado, pues del sacrificio que realiza el altruista no se puede beneficiar cualquiera, sino que, para que el comportamiento altruista pueda evolucionar, es necesario que el individuo seleccione a favor de otros individuos semejantes genéticamente (y/o culturalmente) a él, y que por lo tanto sean los futuros portadores de copias de los mismos rasgos conductuales (genotípicos o fenotípicos) del altruista originario⁶¹.

La defensa teórica del comportamiento altruista no categoriza el altruismo como una pauta monolítica ni de aplicación universal. El altruismo puede, y en cierto modo debe, coexistir con el egoísmo en individuos y en poblaciones incluso mucho más determinados/as hacia el altruismo que otros/as. Así, hay cooperadores y defectores en los experimentos con dilemas sociales, pero ni los unos ni los otros son siempre y sistemáticamente cooperadores o defectores en todas las circunstancias y con todos los co-jugadores⁶². De otro lado, la coexistencia de altruistas y egoístas en una población hace posible la evolución. Por lo tanto, lo que se discute, especialmente desde la revitalización académica de los modelos de selección grupal y cultural, es si sólo existe un único egoísmo individualista universal o si, por el contrario, resulta verosímil la hipótesis de una motivación mixta que incluye egoísmo y altruismo, que propicie procesos simultáneos de selección en varios niveles (selección multinivel): individuo frente a otros individuos dentro de un grupo, y grupo frente a otros grupos integrados por individuos⁶³. Para Sober y Wilson (2000: 15) el altruismo extremo e indiscriminado

⁶¹ La equivalencia genética total sólo se produce entre clones de un mismo ADN, un fenómeno excepcionalmente raro en la naturaleza zoológica. El concepto de semejanza que aquí proponemos es mucho menos rígido que el de la estricta semejanza genotípica o fenotípica. Los humanos consideran la semejanza por proximidad en preferencias, creencias, actitudes, costumbres y patrones de comportamiento y pensamiento, y no por paralelismos morfológicos estrictos u otros marcadores genéticos del tipo feromonas. Esto quiere decir que la categoría mental de semejanza en la especie humana es fenotípica, y por lo tanto su cognición es analógica y no homológica, lo cual no implica, sin embargo, que esté desconectada de los genotipos. Véanse al respecto Efferson, *et. al.* (2008), y McElreath, *et. al.* (2003).

⁶² La palabra defector es un anglicismo no reconocido en el idioma castellano. Sin embargo, dado que su utilización informal está muy difundida en el ámbito académico, aquí lo vamos a utilizar. Su significado es una mezcla entre desleal y desertor, que tal vez se debiera traducir como “inhibido”, “el que se inhibe”, o “el que va por libre”. La figura del defector en interacciones diádicas equivale a la clásica figura del *free-rider* en interacciones n-personales, como las correspondientes a la acción colectiva (Tooby, *et. al.* 2006: 107).

⁶³ Una referencia principal sobre la selección de grupo en la biología del s. XX es Wynne-Edwards (1962, y 1986). Una muestra de la literatura al respecto es: Wilson (1980 y 1983), Wilson y Sober (1994), Michod (1999), Keller (1999), Sober y Wilson (2000), Paulsson (2002), Rayney y Rayney

es raro en la naturaleza, porque no tiene ningún sentido evolutivo. Ellos asocian directamente la existencia de conductas altruistas a la selección de grupo, y por eso no sorprende que definan el comportamiento altruista como conductas que disminuyen la eficacia individual relativa *dentro de* los grupos, mientras que aumentan la eficacia global del grupo respecto a otros agregados de población más amplios (*ibíd.*, p. 78).

Aunque los planteamientos ortodoxos, decantados hacia la tesis de la selección individual (individuo en economía, gen en biología), niegan la verosimilitud de la evolución por selección de grupo en la especie humana (Maynard Smith 1964), sin embargo, es una teoría aceptada que en determinadas condiciones de aislamiento la evolución produce poblaciones de individuos caracterizadas por compartir mayoritariamente un *pool* homogéneo de replicadores (genéticos y/o culturales), y que eso influye en posteriores procesos de selección multinivel y coevolución (Gintis 2003). Las diferencias poblacionales incrementales en cuanto a *pool* genético (y cultural, en el caso del género *Homo*) llegan a producir, a través de procesos de selección de rasgos a largo plazo, especies distintas que no se pueden cruzar sexualmente⁶⁴.

La cultura en el género *Homo* puede haber influido en la evolución biológica a través de procesos de condicionamiento coevolutivo, pero en el caso de la evolución cultural propiamente dicha, las culturas aisladas (al igual que en biología los grupos aislados) teóricamente pueden llegar a producir paradigmas culturales excluyentes entre los que resultaría inverosímil la hibridación (López y Valdaliso 1999: 37)⁶⁵. Por otra

(2003), Traulsen y Nowak (2006), Nowak, *et. al.* (2010), y Hamilton (1975). Algunos de los principales modelos antropológicos de selección grupal basada en replicadores culturales son Boyd y Richerson (1985), Durham (1991), Simon (1990 y 1993), Richerson y Boyd (2005). En economía la selección cultural-grupal ha sido ampliamente tratada por Hayek en varias partes de su obra (véase Ramsay 1987). Una aproximación de la economía experimental en relación con la selección de grupo es la de Henrich (2004).

⁶⁴ Una especie biológica viene definida como «una población de organismos que pueden entrecruzarse potencial o realmente y que comparte el mismo acervo genético» (Gould 2010[1977]: 258). La especiación (el surgimiento de nuevas especies por evolución de otras anteriores, o bien, en el caso humano, por influencias selectivas recíprocas entre la biología y el ambiente cultural) es un proceso explicado por cuatro posibles alternativas de evolución (Cavalli-Sforza 2010[1996]: 50 y ss.):

- a) Mutación, que produce nuevas tipologías genéticas;
- b) selección natural, que criba a los que mejor se adaptan al ambiente (físico, y en los humanos también cultural) en el que viven;
- c) deriva genética, que consiste en el efecto del azar debido a las fluctuaciones estadísticas de las frecuencias génicas de una generación a otra; y
- d) migración de individuos desde una población a otra o de un lugar a otro.

⁶⁵ Incluso en el *Homo sapiens* se dan procesos de coevolución dependientes de la cultura alimentaria.

parte, los primates sociales tienden a manifestar una mayor cooperación con individuos con quienes comparten una muestra del *pool* genético (y cultural, en el caso humano), que con otros individuos menos semejantes (Brosnan 2008). Estos hallazgos invitan a pensar que el factor identificación e identidad juega algún papel en la especificación del comportamiento (cooperativo o no) en primates sociales en general, y en el género *Homo* en particular (Tyler y Blader 2001). La eficacia inclusiva derivada del favoritismo por lo fenotípicamente semejantes es suficiente para explicar este fenómeno constatado experimentalmente en el caso humano:

- a) En psicología experimental existe evidencia empírica de la atracción interpersonal motivada por la semejanza fenotípica (Huston y Levinger 1978);
- b) los circuitos neurales especializados en la empatía se activan con más intensidad ante la percepción de individuos pertenecientes a la misma etnia, una nueva prueba que apoya la hipótesis del favoritismo por la semejanza fenotípica (Xu, et. al. 2009).

Una vertiente especialmente activa de esta controversia académica es la sostenida entre los que Ken Binmore (2006b: 83) denomina neobutlerianos (conductistas de la *behavioral economics*), y los neohobbesianos, entre los que se adscribe él mismo y el resto de los que denomina “economistas ortodoxos”. En esencia la polémica guarda relación con el carácter de la motivación humana:

- a) Los neohobbesianos, defensores de la teoría de la elección racional, se alinean con un monismo motivacional consistente en que toda motivación es egoísta desde el punto de vista utilitario, y los actos altruistas son o errores a corregir con el aprendizaje y la experiencia, o altruismos sólo aparentes⁶⁶;

Así, por ejemplo, se ha demostrado por estudios de genética de poblaciones que al cambiar unos segmentos de población adulta sus hábitos nutricionales y pasar a consumir cantidades significativas de leche procedente de animales domesticados, la actuación a largo plazo de las fuerzas evolutivas seleccionó la genética que posibilita la tolerancia a la lactosa en individuos adultos. El siguiente paso es que por deriva genética casi toda la población ha acabado dentro del *pool* genético que permite la digestión de la lactosa. Intentar una rápida hibridación cultural (que implique el hábito de alimentarse de leche) y biológica entre poblaciones que toleran y no toleran la lactosa ni son ganaderas, puede resultar difícil para los híbridos. Véanse Aoki (1986) y Feldman y Cavalli-Sforza (1989).

⁶⁶ Sobre este último aspecto, Palfrey y Prisbey (1997) enfatizan que el comportamiento cooperativo

- b) los neobutlerianos son defensores de un esquema motivacional mixto, en el que se conjugan motivaciones genuinamente egoístas y genuinamente altruistas, y niegan la excepcionalidad de las conductas altruistas así como su irrelevancia⁶⁷.

Obviamente, ambos enfoques divergen en su punto de vista acerca de la cooperación. Así, los neohobbesianos admiten una cooperación de tipo instrumental motivada sólo por incentivos, necesaria y útil para la supervivencia del individuo, pero en cualquier caso no van más allá de los modelos de selección individual y psicología egoísta. Las restricciones a la conducta individual tienen que ver, en la visión neohobbesiana, con reglas de coordinación útiles para que los individuos egoístas persigan sus intereses sin que la interacción degeneren en un sistema autodestructivo, o en el que resulte imposible que se produzcan equilibrios de coordinación estables⁶⁸. Las instituciones son vistas como epifenómenos de una naturaleza individualista y egoísta, reglamentaciones racionalmente diseñadas o incluso productos de la coevolución cultura-genes.

Los neobutlerianos, de su parte, ven la cooperación como un fenómeno mucho más extendido y complejo, en el que además de los intereses individuales aparecen expresiones conductuales, motivadas por sentimientos morales, relativas a la justicia, preferencias sociales e intereses de grupo. Muchos de sus modelos admiten formas selección grupal y una importante presencia de elecciones y comportamientos empíricos que son presentados como anomalías para los supuestos de la teoría de la elección racional (Dawes y Thaler 1988). En este último caso, se admite la presencia de factores instintivos e institucionales (como las emociones, o la aprehensión de normas mediante procesos informales de socialización o de educación formal) que condicionan *a priori* comportamientos no-deliberados, y que si bien pueden resultar sub-óptimos para el individuo desde una visión económica estrecha o cortoplacista, sin embargo resultan adaptativos desde una perspectiva de racionalidad ecológica, que incluye la

espontáneo en los juegos de dilemas sociales es el resultado de un error de cálculo y/o de la confusión de los jugadores a la hora de comprender correctamente las reglas del juego.

⁶⁷ En general, estos trabajos de referencia son de economía experimental, y concluyen en hallazgos tales como un predominio de la cooperación condicional y de la reciprocidad negativa (castigo altruista a los egoístas) como fundamentos de la evolución de la cooperación. Véase al respecto el monográfico de *Analyse & Kritik* publicado en 2005 sobre la cooperación y el altruismo humanos. También Gintis (2006a).

⁶⁸ Véanse Gauthier (1994[1986]) y Binmore (1998 y 2005).

supervivencia a largo plazo y la replicación del individuo integrado en estructuras sociales duraderas⁶⁹. Las restricciones a la conducta individual serían de tipo epigenético (instintos, y reglas heurísticas innatas fruto de procesos de coevolución), igual que las instituciones, solo que en este caso dentro de una naturaleza humana en la que coexisten el altruismo y el egoísmo como motores de la acción. En general los modelos neobutlerianos admiten supuestos de racionalidad limitada, por lo que las instituciones son el resultado de procesos evolutivos y coevolutivos, complejos, cuya explicación no puede hacerse sólo desde perspectivas culturales y políticas (véase Boyd y Richerson 2001).

1.3.- Los dos problemas teóricos del comportamiento altruista

Existen abundantes evidencias empíricas que prueban el comportamiento altruista. Sin embargo, como ya se ha dicho, hay un sesgo teórico en favor del egoísmo. Aquí defendemos que este escollo puede superarse si se resuelven los dos problemas teóricos más importantes para explicar comportamiento altruista. Ambos están interrelacionados, y son:

- a) La identidad del sujeto, y su influencia en la psicología de la alteridad y las preferencias del individuo.
- b) La perspectiva poblacional en la racionalidad y la toma de decisiones.

Pasemos a examinarlos con mayor detenimiento.

1.3.1.- Insertar la identidad en la psicología y en las ciencias sociales

En su aportación a la controversia sobre el egoísmo Sober y Wilson (2000: 186) juzgan insuficientes las explicaciones positivistas del altruismo consistentes sólo en

⁶⁹ Tooby, *et. al.* (2006: 127n₈) señalan que lo verdaderamente relevante de las interacciones grupales humanas no son ni las interacciones diádicas, ni el pequeño grupo familiar, sino los grupos de mayor tamaño, compuestos por individuos no emparentados, que mantienen relaciones estables de cooperación prolongadas en el tiempo. El *quid* de la cuestión en la selección grupal estaría pues en las diferentes pautas de comportamiento que los individuos desarrollan, para aumentar su eficacia biológica, en relaciones diferentes con diferentes individuos y en diferentes estructuras sociales.

transferencias unilaterales y aparentes conductas de ayuda. Según su enfoque es preciso complementar estas teorías con una psicología de la motivación. Por eso piensan que es esencial establecer previamente una diferencia entre los deseos remotos y los deseos instrumentales⁷⁰. Así, todas las acciones tienen un deseo remoto, cuya realización requiere a veces conseguir satisfacer algún o algunos deseos instrumentales. De este modo, el egoísmo psicológico, o la racionalidad que gobierna las acciones egoístas, consiste en que «todos los deseos remotos están *dirigidos a uno mismo*» (Sober y Wilson 2000: 262; cursivas en el original). Esto implica que los deseos son autodirigidos y causantes de un tipo de acción consecuencialista. Aquí “uno mismo” equivale al individuo-organismo en el que se emplaza el cerebro del que surge un yo que se identifica sólo con el organismo del que emerge. Esto implica que los estados mentales ajenos pueden ser objeto del deseo egoísta, pero no como estados mentales ajenos en-sí-mismos (que “conciernen” directamente a otro organismo), sino porque esos estados mentales ajenos resultan instrumentales al servicio de cierto deseo remoto propio, cuyo cumplimiento aporta algún tipo de beneficio al sujeto⁷¹.

La diferencia entre la psicología egoísta y la psicología altruista estriba entonces en que el sujeto egoísta puede desear que los demás cumplan sus deseos (por ejemplo, que experimenten bienestar, o logren algún objetivo), pero lo hace intencionalmente porque eso le aporta a *él* (al yo-agente) algún beneficio o autosatisfacción. Sin embargo, el agente altruista es capaz de concebir la satisfacción de los deseos ajenos no como un instrumento al servicio de sí mismo *qua* individuo (de *su* yo), sino como un fin intrínseco que afecta a *otro* yo. El altruismo requiere que la capacidad de actuar motivado por deseos remotos referenciados en otros yoes, ya sean estos deseos directamente accesibles, o se trate de meras suposiciones, inferencias o creencias. Como especifica Wispé (1991: 81) apoyándose en las tesis defendidas por Butler en 1726:

⁷⁰ Téngase en cuenta que esta categorización no necesariamente se asocia a la que se refiere la psicología evolutiva como causas próximas y últimas de la acción. Lo que Sober y Wilson denominan deseos remotos e instrumentales, *ambos tipos*, formaría parte de lo que la psicología evolutiva llama causa próxima, puesto que se establece que el individuo es motivado conscientemente tanto por los deseos remotos como por los instrumentales.

⁷¹ Esta sería la postura más influyente en la teoría económica del altruismo, desde Edgeworth hasta Becker. Véanse al respecto los trabajos de Sobel (2005), y Fontaine (2007). Así, el altruismo puede considerarse la disposición a pagar un precio por el “consumo” de una experiencia utilitaria agradable (como el amor), mientras que la discriminación, que es la conducta inversa (aversiva), sería lo mismo pero por evitar una experiencia desagradable (p.ej. el odio, el desdén o el asco). Véase Becker (1997[1971]).

... todos los motivos pertenecen a un yo (tienen un sujeto), y así, la satisfacción de cualquier motivo es *del yo*, pero no necesariamente *para* el yo. (Paréntesis añadido, cursivas en el original).

Esto conlleva reconocer que se puede ser egoísta actuando con benevolencia o con malevolencia; al fin y al cabo en la psicología del egoísmo el ser humano actúa motivado por algún deseo remoto radicado en el propio ego. Al contrario, uno no puede ser altruista actuando con malevolencia si está convencido de que el otro no desea recibir males, algo que aportará una motivación extra al egoísta cuyo deseo consista en producir daño en otro individuo.

La clave de la dicotomía está en comprender que las preferencias (que establecen los deseos remotos) con las que se maneja el individuo en su vida no son siempre *sus* preferencias en un sentido individualista-atomista, sino que representan una amalgama de motivaciones y objetivos con que el individuo-yo se identifica, las hace *suyas*, aunque en realidad pueden ser de otros individuos-yoes, o representar los deseos genéricos (entendidos como elecciones sociales o estereotipos grupales) de una identidad social⁷². Para la psicología evolutiva, por ejemplo, los colectivos sí son tratados como unidades ontológicas representativas, algo así como meta-sujetos, y la explicación estriba en los patrones epistémicos y motivacionales que permiten explicar la acción humana:

... los humanos no tienen problema alguno en hacer interpretaciones sobre los estados mentales atribuidos a grupos, como si un grupo constituyera un agente individual con una mente unificada. (En realidad, obviamente, no lo es) (...) Además, los humanos también tratan sin problema a los grupos de acuerdo a una *teoría intuitiva de los intereses* que pensamos debió evolucionar originalmente porque permitió interpretar la acción individual. Así, los humanos no sólo representan a los individuos teniendo intereses, sino que típicamente también representan a los grupos como entidades que experimentan bienestar y tienen intereses igualmente (..) Así, los grupos pueden tener estatus, rango, estigma, y relaciones de dominación, por no decir amistades y enemistades. (Tooby, *et. al.* 2006: 115; cursivas en el original).

⁷² La hipótesis se encuentra primigeniamente en la psicología social teorizada por Adam Smith (véase Sugden 2002). John Harsanyi (1955) presentó un concepto de “preferencias éticas” en contraposición a lo que él denomina “preferencias subjetivas”. La distinción estriba en que las primeras son algo así como preferencias despersonalizadas capaces de motivar la conducta: «indican la situación social que escogería si no supiera cuál sería su posición en una nueva situación (...) que tuviera una probabilidad igual de obtener cualquiera de las posiciones sociales existentes» (Harsanyi 1955: 316). El planteamiento es parecido al posterior análisis de la “posición original” de Rawls (1972). Ambos son tratamientos racionalistas despojados de un enfoque evolutivo.

Amartya Sen utiliza la noción de meta-ordenamiento aludiendo a «ordenamientos de ordenamientos de preferencias para expresar nuestros juicios morales» (Sen 1977: 337). El autor desea explicar los comportamientos individuales no motivados por la simpatía (que él incluye en el egoísmo consecuencialista) sino en el compromiso con determinada comunidad, nacionalidad, intereses de clase, etnia, sexo, ideología o sistema de valores. Estos comportamientos (no referenciados en deseos remotos propios, sino en un sistema de reglas asociado/definitorio de un agregado social) no son utilitarios, pero no por ello irracionales. Esto es así en la medida que las consecuencias para el agregado social de referencia se entienden y evalúan de acuerdo al deseo de acatar una regulación del comportamiento basada en valores e instintos que, en sí mismos, aunque el autor no lo explicita, constituyen los rasgos definitorios de una identidad social (Sen 1977: 344). En trabajos posteriores Sen sí ha aludido explícitamente al fenómeno de la identificación y su relación con el autoconcepto individual, categorías e identidades sociales (p.ej. en Sen 1985), pero su explicación queda desamarrada de su faceta psicológica y de una explicación científica y evolutiva. No obstante, hoy día existen aportaciones filosóficas y sociológicas acerca de la identidad personal (y su conexión con las identidades sociales mediante procesos neurológicos), para las que se trata de una hipótesis científica admitida sobre la naturaleza humana, así como un tema de estudio en psicología evolutiva y neurociencia cognitiva social⁷³. Las vías innovadoras de investigación que van abriéndose paso en economía y otras ciencias sociales como consecuencia de estos avances científicos se manifiestan en una notable cantidad de trabajos recientes, característicos del arranque de nuevos programas de investigación⁷⁴.

Estas *metapreferencias*, o reglas generales que regulan la interacción, y que son valoradas en sí mismas con independencia de los resultados utilitarios individuales (Kolm 2000b: 124 y ss.), están implantadas en los instintos de un animal social filogénicamente apto para cooperar, así como en los sentimientos morales de un animal

⁷³ Algunos ejemplos son: Cosmides y Tooby (2000b); Klein, *et. al.* (2002); Preston y De Waal (2002b); Decety y Chaminade (2003a y 2003b); Decety y Jackson (2004); Singer, *et. al.* (2004b); Efferson, *et. al.* (2008); Kross (2009).

⁷⁴ Aquí pueden citarse, entre otros trabajos: Zak (2008b); Montero y White (2007); Gintis, *et. al.*, (2005), y Akerlof y Kranton (2000, y 2005).

cognitivo filogénicamente apto para aprender y adaptarse a culturas específicas que propagan normas de comportamiento cooperativo⁷⁵. En el fondo se trataría de un conjunto de reglas epigenéticas adaptadas para la homeostasis del individuo entendido como organismo, y como miembro de un colectivo demarcado de individuos con los que comparte rasgos, hábitat e interacciones (Lieberman, *et. al.* 2003; Loch, *et. al.* 2006). Los mecanismos de selección en varios niveles estarían actuando simultáneamente acoplando conductas maximizadoras de la eficacia biológica del individuo y de los individuos que integran su grupo de afines. Por eso Sober y Wilson presentan una definición de grupo consistente en:

... una serie de individuos que influyen en la aptitud de cada uno con respecto a cierto rasgo, pero no en la aptitud de los que no forman parte del grupo. (Sober y Wilson 2000: 72).

1.3.2.- Motivación altruista *versus* motivación egoísta: la cooperación desde la perspectiva poblacional

Hemos explicado la perspectiva egoísta de Platón y los egoísmos filosóficos posteriores, sean hedonistas o no, de acuerdo a la cual existen deseos remotos e instrumentales, pero en cualquier caso referenciados siempre en el individuo-organismo. Esta concepción egocéntrica de los procesos mentales y de la identidad personal ha tenido un fuerte arraigo en la filosofía y en las ciencias sociales de Occidente (Ryle 1967[1949]; Parfit 2004[1984]). Las perspectivas teóricas egoístas se caracterizan por una axiomática ego-referencial: el despliegue de los deseos siempre está radicado en el yo, la motivación parte de deseos concernientes exclusivamente a *uno mismo* y, puédanse o no conocer los deseos de otros, ello resulta irrelevante, dado que los deseos ajenos sólo son tenidos en cuenta en la medida en que los propios se benefician de ellos vía interdependencias utilitarias⁷⁶. Obviamente, el yo-sujeto de estas perspectivas teóricas es de tipo atomístico, y su única motivación consiste en *su* propia gratificación psicológica⁷⁷.

⁷⁵ Entre cultura y genética existe una estrecha relación que determina qué normas sociales de comportamiento permanecen y evolucionan, y cuáles se extinguen. Como indica Matt Ridley, «las culturas no son colecciones aleatorias de hábitos arbitrarios. Son expresiones canalizadas de nuestros instintos» (Ridley 1996: 6).

⁷⁶ Las interdependencias utilitarias, o externalidades utilitarias positivas, son la causa del comportamiento en la teoría económica del altruismo. Véase Becker (1997[1976a]).

⁷⁷ Caben separar las concepciones metodológicas del individuo-sujeto como átomo, desconectado ontológica y epistemológicamente de otros, y la del individuo-sujeto como parte de un agregado ontológico, y conectado epistemológicamente con otros. Véase Zwirn (2007).

Las perspectivas pluralistas de la motivación asumen supuestos de una psicología más compleja, en el sentido de que admitiendo que en muchas ocasiones los individuos se comportan tal y como explican los teóricos del egoísmo, sin embargo, también lo hacen motivados directamente por intenciones y deseos de otros, así como por deseos compartidos, independientemente de si esas intenciones y deseos ajenos están correlacionados o no con los propios, y generan o no externalidades de utilidad o gratificaciones psicológicas. Los humanos actúan no sólo teniendo en cuenta las intenciones y deseos de otros individuos, sino que a veces lo hacen identificándose plenamente con esas intenciones y deseos, como si siguieran el “interés” del grupo/población al que pertenecen. Identificarse con un grupo es algo bastante habitual en la psicología humana, y consiste en:

... considerarse a sí mismo como una parte de un todo (...) y preocuparse por la suerte de ese todo (...) El “yo” se define relacionado con un “nosotros”. Los seres humanos no *pertenecen* simplemente a los grupos; se *identifican* con ellos. Ésta es una realidad importante de la experiencia humana. (Sober y Wilson 2000: 201; cursivas en el original).

En estos casos, la acción se define como “acción plural” (Schmid 2008), y se caracteriza por no estar motivada por deseos e intenciones concernientes a uno mismo, sino concernientes a otra persona física, o sea, un sujeto individual a quien se atribuye realidad ontológica, o incluso a un sujeto plural que no tiene realidad ontológica, dado que se trata de una construcción mental referida a un conjunto humano, a una población delimitada por rasgos, y por lo tanto, con fronteras más o menos definidas. Cuando un humano se identifica con los deseos e intenciones de otros sujetos o abstracciones conceptuales representativas de sujetos, actúa *autónomamente* (sin coerción) como un sujeto plural o yo trans-personal, y cuando esto ocurre, es un individuo que coopera⁷⁸.

De este modo, la polémica sobre el egoísmo se desliga del hecho empírico de si se llevan a cabo transferencias unilaterales, ayudas en situaciones concretas o sacrificios de cualquier índole; igualmente se desliga de si los agentes obtienen o no bienestar,

⁷⁸ En la literatura psicoanalítica y gestáltica se denomina *introyección* a la identificación del yo con la representación psíquica de ideas, creencias o personas que, de algún modo, “invaden” al sujeto y lo “colonizan”. Asimismo, se denomina *confluencia* a la identificación del yo con un colectivo abstracto, como p.ej. la hinchada de un equipo durante un partido de fútbol. Edgar Morin (2004: 84) denomina *sujetamiento* (sujeto sujetado) a la introyección. Asimismo, sobre la confluencia considera que la propia condición de vivir en un mundo compartido con otros yoes, «instala el Nosotros en la identidad del sujeto» (*ibíd.*, p. 79n₁).

placer, emociones positivas o “utilidad”⁷⁹. El comportamiento altruista no consiste solamente y por definición en conductas de ayuda, donaciones de recursos, y los posibles sentimientos agradables que puedan estar asociados a esas conductas. Téngase en cuenta que estos fenómenos y experiencias también se producen en comportamientos egoístas de aparente altruismo (Andreoni 1990). Más bien, lo que delimita la polémica tiene que ver con la idiosincrasia del sujeto al que se imputan intenciones y deseos (individual o plural), qué relación tiene el yo-agente con ese sujeto (ya se trate de otro individuo o un conjunto de individuos, que admite una concepción estereotipada del conjunto, como p.ej. “los jóvenes” o “los musulmanes”), y cómo esos deseos influyen, por identificación del yo-agente en tal circunstancia, en su motivación para actuar. Esto, en el caso humano, requiere hablar de creencias, del alcance de la toma de conciencia empática y de la psicología de la motivación transpersonal, dado que el “sujeto” del que estamos hablando puede ser un individuo solipsista encapsulado en sí mismo (para quien sólo-existe-él), o un individuo conectado psicológicamente con un agregado de individuos (sujeto plural, o *nosotros*) a través de un proceso de identificación que vincula y coordina la cooperación.

La perspectiva pluralista de la cooperación por identificación requiere la posibilidad de que sea factible adaptar la motivación y la toma de decisiones no sistemáticamente a la idea de individuo-sujeto atomista, como se hace en la teoría de la elección racional, sino a una idea enriquecida de la identidad personal que tenga en cuenta la población⁸⁰. Admitiendo que la identidad y la identificación son operaciones mentales que surgen de la interacción social, en las que el individuo se proyecta a sí mismo en un sujeto no-atomista al que atribuye unas preferencias que asume como *suyas*, se puede plantear una teoría de la acción humana racional, en el sentido convencional de transitividad, concavidad y continuidad, pero que explique la cooperación desde una óptica congruente con los recientes avances de las ciencias del

⁷⁹ En los trabajos de Gary Becker sobre el altruismo se declara de forma explícita la irrelevancia de la motivación (y en general de la psicología, salvo el artificio de la utilidad) en la construcción de los teoremas (Becker 1981: 2). El altruismo consiste exclusivamente en reducir el propio consumo para incrementar el consumo de otro/s (Becker 1997[1976a]: 337). Este enfoque conductista pasa por alto que no es lo mismo ayudar motivado por la necesidad del otro que hacerlo porque desagrada tener ante sí a una persona en apuros. Aunque en ambos casos se maximice utilidad, sin embargo en el primer caso se ayudará efectivamente a la persona en apuros, mientras que en el segundo se eludirá el encuentro con ella, o se reclamará que ayude un tercero (p.ej. los servicios sociales gubernamentales).

⁸⁰ Véanse Sen (1985); Akerlof (1997); Akerlof y Kranton (2000); Schmid (2005a y 2005b).

comportamiento⁸¹. Obviamente, no es cuestión de cambio de teoría, sino de concretar el alcance y el contenido del sujeto al cual se aplica.

En la mente confluyen preferencias y deseos, y el yo-agente actúa en pos de ellos aunque el yo no se identifica necesariamente con un solo organismo⁸². El corolario es que el egoísta es un individuo que actúa motivado sólo para afirmarse a sí mismo; por el contrario, el altruista es un individuo que actúa motivado para afirmar el agregado contextual con el que se identifica y al cual “pertenece” o cree pertenecer, o sea, un “nosotros” o sujeto plural⁸³. Como manifiesta Robert Sugden, la respuesta a la pregunta de cómo puede haber armonía en los grupos sociales

... es que las personas que son compañeros integrantes de una comunidad razonan de manera colectiva, como si cada uno de ellos fuera un parte componente de un sólo agente plural. Esta idea, representada de diversos modos, tales como *sujetos plurales*, *mentalidad colectiva* (*we-thinking*) o *espíritu de equipo* (*team thinking*), resulta extraña a la teoría convencional de la elección racional, en la que sólo las personas individuales cuentan con el estatus de agentes decisorios. (Sugden 2002: 81-82; cursivas en el original).

1.4.- Conclusiones

En este capítulo, dedicado a introducir los antecedentes histórico-filosóficos de esta investigación, se ha podido constatar la existencia de un largo recorrido histórico cultural caracterizado por progresivos intentos, por parte del ser humano, de conocer su propia naturaleza y la de sus relaciones con los demás. Tales iniciativas han estado marcadas por la constante ruptura de límites en el dominio de lo propiamente humano, y por el desplazamiento de lo meramente cultural al ámbito de lo natural en la explicación

⁸¹ El argumento expuesto asume que el cerebro humano está biológicamente diseñado para procesar algoritmos del tipo maximización sujeta a restricciones, y produce comportamientos lógicos y predecibles adaptados en el pasado para optimizar la eficacia biológica (Dennett 1971; Cosmides y Tooby 1994a y 1996). Lógicamente, esto es coherente con el supuesto de racionalidad limitada y sus repercusiones conductistas, entre las que figura que la maximización pueda ser inconsciente (véase Simon 1989: cap. 2).

⁸² Wispé (1991: 81) especifica al respecto un argumento típicamente butleriano: «Toda motivación para actuar pertenece al yo, y por eso la satisfacción de cualquier motivo es *del yo*, pero no necesariamente *para el yo*» (cursivas en el original).

⁸³ Zamagni (1996) concibe los comportamientos prosociales en un sentido muy parecido al que aquí estamos explicando. Según el autor, los individuos expresan comportamientos conducentes a fortalecer las relaciones contextuales con el grupo, y también comportamientos autocentrados sólo en las relaciones con uno mismo. Los primeros son comportamientos prosociales, los segundos son comportamientos egoístas. En la medida que el individuo fortalece las relaciones contextuales con un grupo además de sus propios objetivos, están operando fuerzas de selección multinivel.

de los impulsos que sustentan los sentimientos altruistas y los comportamientos cooperativos⁸⁴. Aquí planteamos un análisis evolutivo centrado en el yo, no entendido como mónada individualista y autosuficiente, sino como fenómeno natural y funcional en la vida orgánica, que es una expresión del aparente afán determinista de la propia vida por permanecer en, y relacionándose con, otros yoes. Esta visión, como se explicará a lo largo del presente estudio, arroja nueva luz sobre cuestiones de la naturaleza humana que ni la filosofía ni las ciencias sociales no evolutivas son capaces de afrontar adecuadamente.

David Hume escribió en el *Tratado de la Naturaleza Humana* que

... la experiencia prueba de un modo suficiente que los hombres, en el curso habitual de la vida, no consideran algo tan remoto como el interés público cuando pagan a sus acreedores, realizan sus compromisos y se abstienen del robo, pillaje e injusticia de cualquier género. Aquel es un motivo demasiado remoto y sublime para afectar a la generalidad del género humano y operar con fuerza en acciones tan contrarias al interés privado como son frecuentemente las de justicia y honradez común. (Hume 1992[1739-1740]: III, ii, 1. p. 310).

Hume, para quien las pasiones (y no el racionalismo) son la fuente de la motivación, explica en esta cita que la gente corriente no contribuye a precipitar la tragedia de los comunes por su comprensión del funcionamiento del dilema del prisionero, o de la lógica del imperativo categórico kantiano. Lo hace por pasiones primarias, como pueden ser las pasiones egoístas racionalizadas en la rúbrica del propio interés utilitario, o bien por pasiones prosociales que inducen al comportamiento altruista por una motivación intrínseca, un aspecto de la naturaleza humana en el que hizo más hincapié si cabe Adam Smith.

El epicentro de la controversia histórica sobre el egoísmo está en si el ser humano es capaz o no de interesarse por el bien de otros, aunque ninguna ganancia inmediata pueda obtener de ello, o incluso aunque le conlleve un sacrificio personal. Se trata en realidad de si las poblaciones humanas están necesariamente abocadas a la tragedia de los bienes comunales (y al conflicto hobbesiano) de no existir derechos de propiedad jurídicamente salvaguardados sobre todos los recursos que son valiosos para

⁸⁴ Agradezco este comentario al profesor Sebastián Álvarez de Toledo, de la Universidad de Salamanca, en el III Encuentro de Estudiantes de Doctorado en Estudios Sociales de la Ciencia y la Tecnología.

el hombre o, por el contrario, es posible concebir una propensión instintiva de preocupación por la comunidad, o preocupación por lo público genéricamente entendido, que opere desde una motivación prosocial. Tal y como señalan Shotland y Stebbins (1983: 36), la literatura sobre el egoísmo y el altruismo se enmarca dentro de estas dos tradiciones intelectuales, que a su vez se plasman en dos enfoques teóricos generalmente excluyentes. Uno consiste en admitir la existencia de un genuino altruismo no intencionalmente utilitario, por el que los individuos experimentan la motivación o impulso de ayudar a otras personas, soportando el coste de dicha ayuda como recursos “perdidos”. El segundo enfoque no contempla la acción de ayuda como un gasto a fondo perdido, sino como una inversión utilitaria; de modo que el altruismo viene condicionado por la expectativa de un beneficio neto (ya se trate de una experiencia de bienestar psicológico, o de cualquier beneficio material derivado de la acción).

Los comportamientos prosociales son externalidades positivas que, según la tradición egoísta, han de ser internalizadas para una provisión óptima que permita la viabilidad de la sociedad. La tradición cooperativa sostiene, por el contrario, que la propia naturaleza humana está capacitada para esa internalización, tanto en los refuerzos instintivos del organismo (que no puede actuar de otro modo), como a través de los beneficios que al cooperador le reporta la propia vida en sociedad. No obstante, la cultura se presenta como un factor adicional que, a través de las innovaciones culturales, puede influir reforzando la motivación humana a través de las creencias, ya sea en una u otra dirección, y de este modo modificar los niveles de cooperación.

Según la tradición cooperativa, lo que diferencia la naturaleza humana de la de otros animales no es la razón o la sociabilidad, sino la capacidad de cooperar sin necesidad de incentivos explícitos. Los humanos cooperan porque son capaces de identificarse con la subjetividad de otros individuos a quienes conciben como sujetos. Por eso son capaces de cooperar con extraños con los que no existen lazos de consanguinidad, sino la comprensión de la similitud cultural (que no deja de ser también biológica).

Desde la perspectiva egoísta la cooperación con extraños sólo se admite si contribuye a los propios intereses del individuo. Se trata de una cooperación consecuencialista, calculada e interesada, un egoísmo racional. Del otro lado, los argumentos partidarios de la hipótesis prosocial presentan la motivación a cooperar como un impulso instintivo que es un fin en sí mismo. No hay pues ni hedonismo ni utilitarismo en esta motivación.

Un punto adicional de abierta discrepancia entre ambas tradiciones es sobre el propio sujeto, o sea, la identidad personal. El yo, según las posturas egoístas, consiste en una mente que recibe información de forma pasiva y se apropia de ella. Con dicha información elabora una biografía que constituye la historia personal del sujeto. El yo, así pues, consiste en un cúmulo de experiencias anudadas por la posesión privilegiada en primera persona; se trata pues de una psicología solipsista en la que la mente aparece reducida a «un solitario lugar de representaciones» (Schmid 2003: 203). Para la tradición cooperativa el yo no actúa pasivamente, sino discriminando y clasificando la información, y reforzando su identidad en esas operaciones. El yo en el primer caso es de tipo identificativo-posesivo, y sólo se ocupa de sus experiencias, dado que, como proclamó por ejemplo Jevons (entre otros), dentro de las tradiciones filosóficas del solipsismo y del utilitarismo,

Cada mente es inescrutable para cualquier otro, y no es posible ningún denominador común de sentimientos. Pero incluso aunque pudiéramos comparar los sentimientos de diferentes espíritus, no necesitaríamos hacerlo, porque un espíritu sólo afecta a otro indirectamente (...) Cada suceso del mundo exterior se representa en la mente por una motivación correspondiente, y es el balance de éstas lo que impulsa la voluntad. Pero las emociones de un espíritu se contrapesan sólo con las emociones del mismo espíritu, nunca con las emociones de otros espíritus. Cada persona es para las otras personas una parte del mundo exterior (el *non-ego*, como le llaman los metafísicos). Así, las motivaciones en el espíritu de A pueden dar origen a fenómenos que se pueden representar por motivaciones en el espíritu de B; pero entre A y B hay un abismo. De aquí que la medida de los sentimientos deba confinarse al seno de cada individuo. (Jevons 1998[1871]: 75-76).

Sin embargo, en el segundo caso el yo es capaz de reconocer lo que le concierne a sí mismo y lo que no le corresponde, y además puede sentirse afectado por las experiencias psicológicas de otros yoes como motivos concernientes. Es decir, se trataría de una psicología tuísta, de un sujeto que dispone de la conciencia de la primera y de la segunda persona (del singular y del plural), y que vive en un mundo

intersubjetivo. De este modo, el sujeto es capaz de identificarse con experiencias vividas comunitariamente y con experiencias ajenas, y por lo tanto es capaz de comprender y actuar con una intencionalidad compartida, y también motivado por los deseos y experiencias de otros que constituyen una población que comparte identidad. Lógicamente, la teoría del sujeto que sostiene las posturas cooperativas no acepta el solipsismo epistemológico ni motivacional, y por ello puede plantear la hipótesis del sujeto plural o yo trans-personal, una perspectiva grupal/poblacional sin la cual no es posible una cooperación evolutivamente estable.

Una vez expuestos los principales debates que han integrado la polémica filosófica, se plantean tres grandes temas, estrechamente asociados con el comportamiento altruista, a estudiar con base científica en el resto de la investigación, y que son:

- a) El sujeto;
- b) la racionalidad; y
- c) las relaciones del sujeto con sus semejantes en una población; o sea, la identidad y la intersubjetividad.

Comencemos con los dos primeros en el siguiente capítulo.

Capítulo 2.- El sujeto y la racionalidad a la luz de la ciencia contemporánea

¡Que tu afán sea por la acción, jamás por sus frutos!
No actúes pensando en los frutos de tu acción,
no te apegues tampoco a la inacción.

(...)

Son dignos de lástima

quienes actúan pensando en los frutos de sus actos.

(*Bagavad Gītā*; II, 47 y 49).

No hay grial más escurridizo o precioso en la vida de la mente que la clave para comprender la condición humana. (Wilson 2012: 1).

...homo non intelligendo fit omnia. (Giambattista Vico)⁸⁵.

Es imposible trazar un dibujo que incluya al dibujante en el acto de dibujar, porque implicaría un “retroceso sin fin”. (Knight 1976[1934]: 281).

(Los seres humanos) tienen una poderosa necesidad de actuar en términos suficientemente razonados. Esta necesidad profundamente asentada de la racionalidad es de hecho una fuente importante de comportamiento irracional. Es también una fuente importante de planes frustrados. (Elster 1993: 256).

Por mucho que pueda pertenecer a Sócrates y a los espíritus como él el adquirir la virtud a través de la razón, hace ya mucho tiempo que el género humano hubiera dejado de existir si su conservación sólo hubiese dependido de los razonamientos de quienes lo componen. (Rousseau 1973[1754]: 63).

La teoría de juegos ortodoxa no puede explicar los resultados (cooperativos observados) (...) Si estos experimentos se contemplan simplemente como tests de la teoría de juegos, la teoría ha fallado. (Ledyard 1995: 172; cursivas en el original. Paréntesis añadido).

⁸⁵ Aforismo que significa: “Todo cuanto el hombre ha hecho lo hizo sin comprender lo que hacía”. Citado en Hayek (1990: 122).

2.1.- Introducción

En el capítulo previo expusimos las líneas maestras del debate filosófico en torno al carácter egoísta o altruista de la naturaleza humana. Concluimos en que hoy día ya se dispone de las herramientas teórico-conceptuales y técnicas necesarias para que el enfoque científico del conocimiento intervenga en el debate estableciendo hipótesis y contrastando teorías. No obstante, la ciencia acoge ingredientes intelectuales procedentes de la reflexión filosófica, y decíamos que en la actualidad resulta posible la contrastación científica de los tres grandes ejes analíticos propuestos para situar la antigua controversia en las coordenadas científicas del s. XX, dígase:

- a) El sujeto;
- b) la racionalidad; y
- c) la identidad y la intersubjetividad.

Pocos han sido los pensadores que desde la filosofía han planteado intuiciones racional y experimentalmente válidas acerca del funcionamiento empírico de la mente. En consecuencia, los debates filosóficos sobre el altruismo no han producido resultados concluyentes. Aunque el supuesto de una naturaleza humana egoísta ha tenido y tiene una notable capacidad expansiva, dado que se trata de una creencia proclive a su replicación (Miller 1999), sin embargo Searle (2000: 188) considera que la hipótesis del egoísmo es un prejuicio académico-cultural carente de rigor. Otros, además, critican su escasa base científica: «Es difícil refutar el egoísmo psicológico, pero también es difícil probarlo» (Sober y Wilson 2000: xiv).

Demostrar los ingredientes altruistas de la naturaleza humana requiere salir de la especulación y estudiar cómo funciona el cerebro en la acción humana real. Para ello se necesita una teoría científica de la mente. El rasgo cooperación ha de ser incorporado a la visión científica del mundo porque el altruismo es un hecho empírico igual que egoísmo, que no requiere una justificación especial ni es la excepción que cumple la regla (del egoísmo) sino que, como plantea Searle, el rasgo altruismo presente en la naturaleza humana no debe considerarse ni más ni menos especial que cualquier otro. El

mismo animal que se comporta de manera egoísta es capaz de hacerlo de manera altruista, y debe existir algo unificador en esos rasgos aparentemente antitéticos.

La perspectiva teórica que aplicamos en este capítulo es la de la psicología evolutiva contemporánea, que explica la mente humana y sus capacidades (percepción, aprendizaje, memoria, inteligencia, pensamiento, voluntad y emociones) como productos de la evolución y, en este sentido, surgidos de una senda evolutiva que recoge patrones de actividad mental comunes a otros animales (Barkow, *et. al.* 1992). El sistema límbico, la estructura más antigua del cerebro animal, gestiona procesos emocionales semejantes en todos los animales que disponen de él. El neocórtex, la parte más reciente en la evolución del cerebro, distingue al *Homo sapiens* en complejidad (dado el tamaño diferencial del neocórtex humano) más que en novedad. El neocórtex es la estructura donde reside el modelo cognitivo de identidad personal, la comprensión lógica, la elaboración de pensamientos y juicios, y la capacidad del lenguaje y de realizar una comunicación consciente.

El pensamiento reflexivo, un nivel superior de complejidad mental que posibilita la comprensión abstracta y la formulación de hipótesis y teorías, nos lleva equivocadamente a concebir el animal humano como ente angélico, dueño de su destino y liberado de los instintos que gobiernan el comportamiento del resto de seres vivos. Así, el ser humano se ha explicado como animal racional e inteligente, y su comportamiento fruto de la deliberación racional y no de los instintos⁸⁶. Sin embargo, la psicología evolutiva explica el pensamiento consciente y la lógica racional humanas como instintos, que dependen en última instancia de una configuración genética, igual

⁸⁶ La evocación especulativa de esta orientación es obvia:

... la inteligencia ejerce sobre el apetito un dominio político regio. En esto resulta evidente que es conforme a la naturaleza y provecho para el cuerpo someterse al alma, y para la parte afectiva, ser gobernada por la inteligencia y la parte dotada de razón, mientras que disponerlas en pie de igualdad, o al contrario, es perjudicial para todos. (Aristóteles, *Política*, I, v, 1254b).

Una especulación que por siglos permaneció inalterada:

... una voluntad que no puede ser determinada nada más que a través de estímulos sensibles, es decir, *patológicamente*, es una *voluntad animal (arbitrium brutum)*. La que es, en cambio, independiente de tales estímulos y puede, por tanto, ser determinada a través de motivos sólo representables por la razón, se llama *voluntad libre (arbitrium liberum)*, y todo cuanto se relaciona con ésta última, sea como fundamento, sea como consecuencia, recibe el nombre de *práctico*. La libertad práctica puede demostrarse por experiencia, puesto que la voluntad humana no sólo es determinada por lo que estimula o afecta directamente a los sentidos, sino que poseemos la capacidad de superar las impresiones recibidas por nuestra facultad apetitiva sensible gracias a la representación de lo que nos es, incluso de forma remota, provechoso o perjudicial. Estas reflexiones acerca de lo deseable, esto es, bueno y provechoso, en relación con todo nuestro estado, se basan en la razón. (Kant 1978[1781]: II, ii, 1, p. 628. Cursivas en el original).

que todo lo demás (Cosmides y Tooby 1994a y b): los llamados innatismos⁸⁷. Este enfoque cuenta con el antecedente teórico-filosófico de Spinoza en el s. XVII, y de la teoría y filosofía de la mente elaboradas por Hayek en el s. XX. En ambos casos, expuestos en el apartado 2, se matiza el concepto ideológico de libertad humana, tal y como se presenta en la teoría de la decisión y en la psicología popular.

Desde la perspectiva biológica, el sentido evolutivo de la vida consiste en la continuidad de la misma a través del éxito reproductivo (Dennett 1995). El éxito se logra a través de la adaptación diferencial a las condiciones cambiantes. Todos los instintos de los animales están orientados teleológicamente a esa misión existencial. En este propósito autorreplicador y adaptacionista del fenómeno vida, De Waal (2008: 63-64) explica que los organismos actúan sobre la base de dos tipos de causas del comportamiento⁸⁸: por un lado, lo que denomina “dominio de causación última” (análisis evolutivo de la conducta), que obedece a los innatismos que determinan al individuo a sobrevivir y reproducirse, y que son el resultado de adaptaciones filogenéticas ocurridas en el pasado y codificadas en el ADN. Estos procesos e impulsos son innatos, en alto grado inconscientes, y todos los seres vivos disponen de ellos. Por otro lado, está el “dominio de causación próxima”, o causas próximas del comportamiento (un análisis funcional de la conducta), que se refieren a los procesos orgánicos epigenéticos que actúan ante situaciones inmediatas y estímulos que desencadenan estas situaciones, y que determinan el comportamiento en unas circunstancias específicas determinadas. Obviamente, las causas próximas están engarzadas en las causas últimas, dado que, salvo casos patológicos, todo organismo actúa para sobrevivir y reproducirse. Pero en realidad los organismos se comportan inducidos por causas próximas, y no motivados (cognitivamente hablando) por las causas últimas. Los procesos biológicos que gestionan la causación próxima son programaciones epigenéticas adaptadas al medio durante la ontogénesis del individuo; tales procesos pertenecen al dominio de los innatismos, si bien desarrolladas en la experiencia vital del organismo. Los organismos dotados de cerebro, como los

⁸⁷ Desde el paradigma de las ciencias cognitivas no se habla de instintos, sino de innatismos, y son descritos en la *Enciclopedia MIT de las Ciencias Cognitivas* como un «cuerpo significativo de conocimiento que está “incorporado” al organismo, o al menos predeterminado de forma innata» (Frank Keil, en Wilson y Keil 2002: 643.)

⁸⁸ La primera aproximación teórica a este enfoque de las causas del comportamiento animal se debe a Mayr (1961: 1.502-1.503).

humanos, disponen de algún grado de conciencia sobre el mundo que les rodea y de su situación en él, y sus acciones están enmarcadas en esa faceta cognitiva desde la perspectiva de un sujeto (Sober y Wilson 2000: 185). En este sentido, se puede hablar de organismos psicológicos. Así, por ejemplo, el comportamiento sexual tiene un dominio de causación última que es la replicación genética, pero los animales desconocen ese mecanismo de causación de su comportamiento (seguramente sólo los humanos son conscientes de ello). Lo que sí perciben muchos organismos, además de los estímulos sexuales, son los estados emocionales relacionados con sexualidad, y actúan movidos por tales experiencias (causas próximas), que pueden ir asociadas a sensaciones de bienestar, antes, durante o después de la acción:

Comemos y practicamos sexo porque disfrutamos con ello, pero detrás del placer figura el programa evolucionista que nos ordena sobrevivir y procrear. En una línea similar, las emociones sociales como la amistad, la vergüenza, la generosidad y la culpa nos aguijonean para conseguir el éxito biológico en redes sociales complejas. (Sigmund, *et. al.* 2002: 87).

En el presente capítulo se analiza el soporte neurológico que explica el dominio de causación próxima del comportamiento humano. El objetivo de fondo, que pertenece a los capítulos 5 y 6 de esta investigación, consiste en explicar las causas próximas del comportamiento altruista en la especie humana. Pero previamente a ello es preciso conocer cómo funciona el cerebro para generar la mente y la psicología de la primera persona (apartado 3), así como el modelo de comportamiento (racional) humano (apartado 4). Este análisis es preciso porque el altruismo humano es un fenómeno natural que depende de esos condicionantes neurobiológicos.

La causación próxima del comportamiento animal está asociada a un genoma específico, y dentro de ese genoma al sistema nervioso dependiente de aquél. Explica Gazzaniga (2006: 21) que la neurociencia «estudia el órgano que nos identifica como humanos: el cerebro que permite una vida consciente». La hominización se debe a transformaciones en el ADN que dieron lugar a una nueva especie, pero en lo que supone la expresión fenotípica de esas transformaciones, la más significativa en cuanto a la propia definición de la naturaleza humana es la del sistema nervioso. En el aumento progresivo del volumen encefálico y de la complejidad neural se encuentra la explicación de la conciencia reflexiva, sujeto o yo (en el pasado llamado *alma*).

Asimismo, en el apartado 3 se explica cómo es en el cerebro donde se desarrollan los procesos conducentes a la acción, y esos procesos generan el espejismo de un yo que toma decisiones libremente y gobierna el cuerpo al estilo del dualismo cartesiano. Del mismo modo, el sistema nervioso contiene y procesa los programas instintivos que explican tanto la configuración de un yo autónomo, la sensación de libertad decisoria, el pensamiento y la racionalidad, y otros aspectos relevantes como la naturaleza cultural y social del *Homo sapiens*. Éstas son las señas de identidad de la especie diferenciadas de las de otros primates sociales y homínidos anteriores y coetáneos al hombre moderno (Levi Montalcini 2000: 89-90).

Las implicaciones axiomáticas de las tradiciones filosóficas defensoras de una naturaleza humana egoísta han llegado a nuestros días influyendo en concepciones teóricas de la racionalidad que, como el caso de la economía y la teoría de juegos, parten de supuestos conductuales de racionalidad y egoísmo psicológico entendidos como rasgos sinónimos o asociados. Este fenómeno a su vez se ha ido extendiendo a las ciencias sociales en su conjunto desde el axioma de la elección racional. Por un lado, se presenta el comportamiento humano como producto del libre albedrío de un yo que se caracteriza por ser racional y ejecutivo. Sin embargo, la ciencia contemporánea, que sigue el monismo explicitado en las tesis spinozistas, cuestiona esa línea de pensamiento no sin generar nuevas polémicas: «Decir que el comportamiento tiene una base genética obliga a reflexionar sobre la verdadera dimensión de la libertad individual» (Ferrús 1998: 80). Por ello Patricia Churchland (2006) propone que es necesario revisar el concepto tradicional del libre albedrío, si bien dentro de unas coordenadas científicas que eviten dañar el concepto de responsabilidad de las acciones, sobre el cual se erigen los sistemas morales, políticos y sociales que sustentan la compleja sociabilidad humana⁸⁹.

Otro concepto procedente de la filosofía que hoy día la ciencia está sometiendo a revisión es el de racionalidad entendida como producto deliberado de una mente consciente, y por ello atributo exclusivo de la especie humana. Esto es otro asunto

⁸⁹ La revista *Foreign Policy* publicó en 2004 un monográfico sobre las ocho ideas o ideologías más peligrosas del mundo actual. Una de ellas es la posibilidad de que los descubrimientos que la ciencia va mostrando, acerca de cómo el cerebro genera la sensación de control y libre albedrío, puedan ser utilizados al servicio de objetivos de manipulación y dominio de las personas (Davis 2004).

complejo y polémico, puesto que la racionalidad aparece asociada no sólo al libre albedrío, sino también al egoísmo psicológico y al utilitarismo. Como se explica en los sub-aparados 4.3. y 4.4., la neurociencia desliga estos conceptos y muestra que el ser humano no es un animal racional diferente de los demás en ese sentido restringido (subapartado 4.1.), sino que más bien es un animal racionalizador de conductas (subapartado 4.2.):

La evolución cognitiva no inventa nuevas categorías de conductas, ni sustituye la antigua infraestructura emocional; colabora con ella, transformándola por medio de una mayor comprensión por parte de los actores. Sin embargo, en tanto que especie, valoramos hasta tal punto la inteligencia que tendemos a pensar que nuestra capacidad de razonamiento *impulsa* la conducta. Somos tan hábiles para producir racionalizaciones convincentes que al final acabamos creyéndonlas: se trata del mito del hombre, o de la mujer, racional. (De Waal 1997: 104; cursivas en el original).

En su afán por demostrar que la razón no es una superposición de innatismos y cultura, el pensamiento racionalista ha seguido programas de investigación de corte reduccionista, hasta ampararse en la teoría de juegos. La elección de la estrategia dominante en juegos de dilemas “sociales” (aunque tan sólo se opta por una pareja genérica de individuos), como es el caso de los juegos de dilema del prisionero o de sostenimiento de los bienes públicos a través de la acción colectiva, lleva a plantear la racionalidad de no cooperar en ambos casos, especialmente si se trata de juegos finitos, e incluso juegos de una única tirada y en ambientes de anonimato (*single spot games*). Además, en estos casos se evita deliberadamente el análisis desde la perspectiva grupal/poblacional. Lógicamente, si los pagos y la psicología de los individuos se plantean tal y como se hace normalmente, lo racional es elegir la estrategia dominante, o sea, la no-cooperativa. Sin embargo, cuando la función de utilidad del individuo se le exime de ser una consecuencia de la razón, y se inducen supuestos evolutivos en ella, como la influencia de las emociones en la elección, y no sólo el cálculo de beneficios materiales inmediatos, los pagos se alteran en la mente de los jugadores y la estrategia dominante puede pasar a ser cooperativa en lugar de defección (Sally 2000; Colman 2005). Las emociones -término con el que hoy día se alude a una parte de los innatismos con que se procesan las decisiones- condicionan la elección racional, y como se explica en el apartado 4.2. no es posible pensar de forma aséptica sin la impronta que las emociones transmiten al raciocinio. Sartre lo intuyó acertadamente cuando explicó que:

No cabe, pues, considerar la emoción como un desorden pasajero del organismo y de la mente que turbe desde fuera la vida psíquica (...) La emoción no es un accidente, sino un modo de existencia de la conciencia. (Sartre 2005[1965]: 98).

No obstante, todavía muchos modelos de racionalidad identifican la racionalidad con el cálculo utilitario de beneficios materiales, sin tener en cuenta que el cerebro procesa en función del etograma (Dawes y Thaler 1988). Los innatismos que la mente humana despliega (nuestro etograma) hacen que interpretemos las decisiones adoptadas como emociones controladas deliberadamente por la esfera cognitiva y calculadora de la mente, para convencernos de que se está persiguiendo el propio interés (atomista y consecuencialista), cuando en realidad la decisión ya ha sido automatizada unas décimas de segundo antes de su “racionalización” (Libet 1996; Gazzaniga 1995). En efecto, en relación a la toma de decisiones (el comportamiento), la razón es un mecanismo cerebral de autoconvicción. Este aspecto resulta relevante para abordar el estudio del altruismo, puesto que los procesos decisorios en homínidos tienen una base instintiva de connotaciones y evaluaciones emocionales y morales (Hauser 2006; Young, *et. al.* 2007). Puesto que toda especie social evoluciona dentro de ecosistemas poblacionales, y dado que el individuo vive en ese mundo por naturaleza, por eso mismo ha de estar capacitado para tener en cuenta la dimensión grupal/poblacional, en la que se incluye el condicionamiento cultural de sus comportamientos. Pero el tamiz social y moral viene dado por una neurofisiología adaptada a la vida social, que permite la interacción de emociones sociales y de la facultad de la anticipación y planificación en los programas neurales que se activan cuando el cerebro toma decisiones. Esto es un producto de la evolución que se sustancia en innatismos, y no un refinamiento derivado del pensamiento racional humano que supera la eficacia de la biología (la *fitness* evolutiva) a la hora de acertar en la decisión correcta:

El biólogo que se ocupa de los problemas de la fisiología e historia de la evolución, se da cuenta de que su propio conocimiento está forzado y condicionado por los centros de control emocionales situados en el hipotálamo y en el sistema límbico del cerebro. Estos centros inundan nuestra conciencia con todas las emociones –odio, amor, culpabilidad, miedo y otras- estudiadas por los filósofos de la ética que desean intuir los patrones del bien y del mal. De este modo, nos vemos obligados a preguntar ¿qué es lo que originó los sistemas límbico e hipotalámico? Esos sistemas evolucionaron a través de la selección natural. Debe hacerse hincapié en esta simple afirmación biológica para entender no sólo la ética y a los que la estudian, sino también la epistemología y los epistemólogos. (Wilson 1980[1975]: 3).

2.2.- La acción humana, o la libertad desde una perspectiva científica: Spinoza y la Escuela Austríaca

Spinoza formula en el prefacio de la parte tercera de la *Ética* una crítica a la tradición racionalista que presenta los instintos y emociones (las “pasiones” o los “afectos”, en el lenguaje de su época) como propiedades degradantes de la naturaleza humana, por ser motivaciones de carácter inferior al pensamiento racional:

La mayor parte de los que han escrito acerca de los afectos y de la conducta humana, parecen tratar no de cosas naturales que siguen las leyes ordinarias de la naturaleza, sino de cosas que están fuera de ésta. Más aún: parece que conciben al hombre, dentro de la naturaleza, como un imperio dentro de otro imperio⁹⁰.

Se trata de una crítica al dualismo cartesiano y su hipótesis disociativa de las esferas corporal y mental de la naturaleza humana, una perspectiva que se remonta al menos hasta el mito platónico del auriga⁹¹. El racionalismo sostiene que existe una noble manifestación de la humanidad en la razón frente a los primitivos instintos, y que es el pensamiento racional, y no los instintos, la justificación de la libertad propia de lo humano, la inteligente guía de la motivación, causa de la voluntad y de la libertad de acción⁹². Spinoza discrepa de esta interpretación: «quienes creen que hablan, o callan, o hacen cualquier cosa, por libre decisión del alma, sueñan con los ojos abiertos» (*Et.* III, ii, *e*).

El monismo spinozista parte de la base de que hay unidad entre los procesos orgánicos del cuerpo, que se expresan en la forma de percepciones, sensaciones, apetitos, deseos, pensamientos y movimientos. La voluntad, que en apariencia es la causa de la acción, no es un proceso realimentado desde una mente autónoma y

⁹⁰ En lo sucesivo, cuando citemos textos de la *Ética* de Spinoza (2004[1677]) lo haremos con la notación *Et.*, seguida entre paréntesis de la parte, de la proposición citada, y, si es el caso, del corolario (*c*) y del escolio (*e*).

⁹¹ Spinoza se desmarca inmediatamente del dualismo cartesiano afirmando que el alma (el pensamiento consciente, dentro de lo mental) y el cuerpo (lo inconsciente y lo físico-orgánico) son «una sola y misma cosa, que se concibe, ya bajo el atributo del pensamiento, ya bajo el de la extensión» (*Et.* II, vii, *e*).

⁹² Esta postura se erigió en la idea central de la Ilustración europea, y ha llegado a nuestros días a través de la filosofía y de las ciencias del comportamiento. El denominado por John Tooby y Leda Cosmides “modelo estándar de las ciencias sociales” supone una visión moderna y readaptada del mito platónico del auriga y del dualismo cartesiano (véase Pinker 2004: 68 y ss.) Sin embargo, es un fenómeno cultural específico de la civilización occidental moderna, que no dominó el pensamiento europeo anterior, ni tampoco otras tradiciones culturales no europeas (Herrmann-Pillath 1994: 56-57).

desligada del cuerpo, sino que lo que creemos voluntad o libre elección es una manifestación mental de los procesos motivacionales del cuerpo. La motivación, lo que “mueve” el comportamiento humano, consiste en lo que Spinoza denomina “afectos”. El gran drama humano, sin embargo, es que el hombre desconoce y no comprende esta verdad científica. Los hombres creen ser libres porque son conscientes de que actúan, pero son esclavos de sus pasiones porque ignoran o tienen ideas confusas sobre las causas que determinan su acción. Esto les lleva a concebirse a sí mismos de forma ilusoria⁹³. En una carta escrita en 1674 expuso el núcleo de su teoría de la acción humana y su tesis de la ficción del libre albedrío y la voluntad. Esta teoría pre-científica será la que posteriormente desarrollará en la *Ética*:

Tal es la libertad humana que todos los hombres se vanaglorian de tener, y que tan sólo consiste en que los hombres son conscientes de su apetito e ignorantes de las causas que lo determinan. Así, el niño cree apetecer libremente la leche, el muchacho irritado que quiere la venganza, y el tímido la fuga. El borracho cree decir por libre decisión de su alma lo que después, ya sobrio, hubiera querido callar. Igualmente, el que delira, el charlatán, el niño y otros muchos casos de este tipo, creen obrar por decisión libre de su espíritu, y no que son llevados por un impulso. Y como este prejuicio es innato en todos los hombres, no se liberan tan fácilmente de él. Y aún cuando la experiencia enseñe, más que sobradamente, que los hombres nada pueden menos que dominar sus apetitos, y que muchas veces, mientras sufren la pugna de efectos contrarios, ven lo mejor y siguen lo peor, sin embargo, se creen libres por la sencilla razón de que tienen por objeto el deseo de algo, y que ese deseo puede ser fácilmente reprimido por el recuerdo frecuente de otra cosa que les viene a la memoria. (Spinoza, 1988[1674]: 265-266).

¿Cuál es el contenido de la libertad en la teoría spinozista del comportamiento?

El autor sólo considera una verdadera libertad aquel comportamiento que es una causa incausada. Las causas causadas, por lo tanto, solamente son una libertad aparente. Así,

... los hombres se equivocan al creerse libres, opinión que obedece al solo hecho de que son conscientes de sus acciones e ignorantes de las causas que las determinan. Y, por tanto, su idea de “libertad” se reduce al desconocimiento de las causas de sus acciones, pues todo eso que dicen de que las acciones humanas dependen de la voluntad son palabras, sin idea alguna que les corresponda. (*Et*: II, xxxv, e).

¿Cuáles son, pues, las causas reales de las acciones humanas? Existe según

Spinoza un principio elemental que caracteriza a todos los seres vivos y que determina

⁹³ Como argumenta Gilles Deleuze, Spinoza dedicó su incompleto *Tratado sobre la reforma del entendimiento*, así como la Parte II de la *Ética* (obra publicada póstumamente, al igual que el primero, en 1677) a explicar cómo se produce esta ilusión del yo autónomo y autogobernado: «Ocurre que la conciencia es naturalmente el lugar de una ilusión. Su naturaleza es tal que recoge los efectos pero ignora las causas» (Deleuze 1984: 29-30).

la regulación natural de su comportamiento: el *conatus*⁹⁴. De acuerdo al *conatus* «cada cosa se esfuerza, cuanto está a su alcance, por perseverar en su ser» (*Et*: III, vi). En el hombre el *conatus* adquiere una dimensión psicológica, una conciencia de estados corporales, movimientos internos y pensamientos, que no se da en los demás animales, o que se produce de una manera que desconocemos⁹⁵. El autor se refiere a esto como «las ideas de las afecciones del cuerpo», y es la necesaria conciencia de estas afecciones lo que permite trazar la hipótesis central spinozista sobre la identidad personal, entendida de este modo como conciencia del “esfuerzo” por ser (existir) y seguir siendo (existiendo) el entramado biológico que se es, unificado por un flujo incesante de afecciones: «El hombre no se conoce a sí mismo sino a través de las afecciones de su cuerpo y las ideas de éstas» (*Et*: III, liii). La idea de identidad personal se presenta aquí como una estabilidad en la existencia orgánica, que tiene su reflejo en los procesos corpóreos y mentales del organismo. El *conatus* define y delimita eso-que-se-es por medio de afecciones que implican al cuerpo y a la mente (a través de los *apetitos* que afloran a la consciencia), o sólo a la mente (la *voluntad*) (*Et*: III, ix). Las afecciones que determinan a la mente, es decir, la voluntad, según Spinoza sólo son propias de la especie humana. No obstante, esto no debe entenderse como libre albedrío, sino como un esfuerzo intelectual, consciente, de estabilidad y defensa de la existencia orgánica, que no puede ser una libre decisión, puesto que resulta imposible contravenir la regulación pre-programada de los instintos.

¿En qué consiste pues “decidir”? Spinoza se muestra explícito, indicando que las decisiones son los propios apetitos del cuerpo, que a veces se muestran contradictorios, generando indecisión o duda, y que otras veces no se manifiestan y quedan en estado de latencia, causando lo que en filosofía se denomina debilidad de la voluntad o veleidad. Lo que suponemos una libre decisión es en realidad la sincronización de procesos

⁹⁴ Aunque caben precisiones terminológicas, aquí vamos a identificar la expresión spinozista *conatus* con lo que hoy día llamamos emociones, aunque puede extenderse al conjunto de innatismos, o instintos de un organismo. Véase Damasio (2005).

⁹⁵ Damasio (2005: 85) afirma que los sentimientos traducen el estado vital en curso en el lenguaje de la mente. Se trata de “metarrepresentaciones” de los estados internos del cuerpo, y por lo tanto un fenómeno simbólico. Las emociones son a los sentimientos lo que son las reacciones fisiológicas del organismo al pensamiento, y define los automatismos emocionales como dispositivos diseñados por la evolución para resolver *automáticamente*, sin deliberación racional, los problemas básicos de la vida. Tales problemas básicos son: encontrar fuentes de energía, mantener un equilibrio químico compatible con el proceso vital, conservar la estructura del organismo mediante la reparación del desgaste natural, y detener los agentes externos de enfermedad y daño físico.

físico-orgánicos y mentales que consisten en el afloramiento de lo corpóreo a la esfera consciente de la psique mediante una representación:

... tanto la decisión como el apetito del alma y la determinación del cuerpo son cosas simultáneas por naturaleza, o mejor dicho, son una sola y misma cosa, a la que llamamos “decisión” cuando la consideramos bajo el atributo del pensamiento, y “determinación” cuando la consideramos bajo el atributo de la extensión. (*Et*: III, ii, e).

Así, en esa interacción entre las facetas corpórea y mental no hay indeterminación, sino determinación, puesto que el organismo funciona y produce experiencias psicológicas sobre la base de lo que ya “sabe hacer”, esto es, sobre los carriles ya trazados en la filogénesis de la especie y la ontogénesis del individuo. Spinoza habla, como Hayek en *El orden sensorial*, de memorias y recuerdos, pero en modo alguno de libertad de la razón, para ejercer algo que supuestamente no “sabe” hacer el organismo: «nosotros no podemos, por decisión del alma, hacer nada que previamente no recordemos» (*Et*: III, ii, e)⁹⁶. Consecuentemente, nuestra libertad, como Spinoza explica en la Carta a Schuller, no es un libre decreto, sino una libre necesidad.

Un sistema de pensamiento en el que la libertad es un concepto central es el de Hayek, para quien la libertad no consiste en el libre albedrío racionalista, sino en que el individuo pueda perseguir sus propios fines en ausencia de coacción o sujeción a la voluntad arbitraria de otros (Hayek 1991[1960]: 27). Y es que Hayek, que considera espejismos de la mente tanto la voluntad como la libre intención y designio, rechaza el consecuencialismo implícito en la tradición racionalista⁹⁷. Esto conlleva un

⁹⁶ Hayek (2004[1952]: §5.41, y ss.) utiliza los conceptos “mapa” y “modelo”, y vienen a ser programaciones (entre otras, neurológicas) con las que funciona el organismo y el cerebro, en los procesos metabólicos y en la representación tanto de estados internos como de interacciones sensoriales con el mundo a través de la percepción. De este modo,

... la experiencia sensorial presupone la existencia del algún tipo de “conocimiento” acumulado, de un orden adquirido por los impulsos sensoriales basado en su aparición conjunta en el pasado; y este conocimiento, aunque basado en la experiencia (pre-sensorial), nunca puede ser contradicho por experiencias sensoriales y determinará, además, las formas en que tales experiencias son posibles. (*Op. cit.*, §8.8).

⁹⁷ Se puede decir que Hayek se alinea con la teoría del “determinismo blando” de Ayer (1979[1954]: cap. 12), según la cual, aún admitiendo un determinismo fiscalista -que incluye la evolución cultural, a través de los procesos de darwinismo neural-, se considera una persona libre siempre y cuando realice acciones libres, y son acciones libres aquéllas que son el resultado de procesos psicológicos subjetivos (como deseos, creencias, intenciones y decisiones) sin amenazas o imposiciones externas. Así, la consideración de libertad de acción no se basa en la existencia o no de una causa orgánica (determinismo neurológico), lo cual es indiscutible, sino en el organismo donde se origina (el cerebro del actor, *versus* otro cerebro).

cuestionamiento directo del sentido dado al libre albedrío en la concepción utilitarista del comportamiento humano, entendida la agencia por parte de un “sujeto” que persigue conscientemente objetivos previamente seleccionados a través de medios instrumentales. En Hayek todo comportamiento se basa en reglas pre-existentes. Estas reglas equivalen a grandes rasgos al *conatus* spinozista. En el autor austríaco las reglas son de base instintiva, y por ello genética, o bien aprendizajes culturales de costumbres y valores, que se erigen sobre la base epigenética de los instintos. De ahí que Hayek no sitúe la acción humana en la razón especulativa, calculadora ni práctica, sino en un apartado de la naturaleza humana que se ubica exactamente «entre el instinto y la razón» (Hayek 1990: 57). En efecto, el comportamiento humano no está motivado ni en la elucubración racional, ni en la selección lógica *a priori* de fines últimos y objetivos instrumentales, sino en reglas que se implantan en la neurología del individuo. Sobre la base de las reglas epigenéticas innatas, se implantan otras reglas, las morales, que son asimiladas del medio social, y una vez aprendidas por condicionamiento hebbiano, vienen a operar de manera muy parecida a los instintos congénitos de la especie⁹⁸. Argumenta Hayek:

... la mayor parte de las normas, instituciones y prácticas propias de la moralidad tradicional, son “irracionales” y “acientíficas” (...) Cualesquiera que sean las normas que en definitiva decidamos adoptar, nunca podrán justificarse sobre la base de la razón... Si dejásemos de ajustar nuestra conducta a cuanto racionalmente no podemos explicar -o a aquello que hacemos sin saber por qué lo hacemos-, hasta nuestra propia existencia correría peligro (...) No hay ninguna razón para suponer, como lo hacen los partidarios del racionalismo y del hedonismo, que nuestra moral deba apuntar a la consecución de determinados fines; por ejemplo, el logro de la felicidad. Nada induce a pensar que la selección evolutiva de pautas de comportamiento más adecuadas para facilitar la supervivencia de una más nutrida población tenga nada que ver con el logro de la felicidad de nadie, y mucho menos que sea fruto de algún intento de alcanzarla. Por el contrario, son muchos los indicios de que si alguien hubiera intentado alcanzar de este modo su propia felicidad, habría sido sin duda arrollado por cuantos se esforzaban simplemente por sobrevivir. (Hayek 1990: 117, y 120-121).

En Spinoza el único margen de auténtica libertad posible está en conocer que no somos libres al estilo del racionalismo (y por lo tanto no somos “dueños” de nuestros actos), y en afirmar o negar guiados por la razón, dentro de las posibilidades limitadas

⁹⁸ Véase Hebb (1985[1949]), quien desarrolló las ideas principales que Hayek explica en *El orden sensorial* con un lenguaje más técnico. Donald Hebb fue pionero en la biología del comportamiento, un enfoque que podemos considerar antecedente de la neurociencia actual. Para Hebb las funciones de la mente y las experiencias mentales son el resultado de actividad conectiva entre dendritas y axones neuronales. La repetición de movimientos y estímulos reitera la actividad conectiva y contribuye a reforzar circuitos ontogenéticos que explican el aprendizaje y la memoria, o sea, la composición de la identidad personal relativa a la interacción con el medio.

que dispone la razón humana. Los órdenes espontáneos que gobiernan el cosmos sólo son cognoscibles intuitivamente, se llega a ellos por la guía de la razón, pero es imposible un conocimiento directo e inductivo de ellos⁹⁹.

En esta línea Ludwig Mises, otro autor perteneciente igual que Hayek a la tradición austríaca de filosofía social, especifica que:

Ningún impulso instintivo puede ser analizado de modo racional, ya que la razón se ocupa sólo de los medios idóneos para alcanzar los fines deseados, pero no de los fines últimos en sí. El hombre se distingue de los restantes animales en que no cede a los impulsos instintivos si no es con un cierto grado de voluntariedad. Se sirve de la razón para, entre deseos incompatibles, optar por unos u otros. (Mises 2009[1949]: 207).

Así, la razón, que actúa como un «pistolero a sueldo», que «no nos puede decir adónde ir, (sino) como mucho nos indica cómo llegar allí» (Simon 1983: 7), aparentemente “decide”. Sin embargo, en lo que consiste realmente el raciocinio es en un proceso consciente que relata una cesión a los impulsos instintivos, unos sobre otros, y, entre ellos, un impulso instintivo adicional (que es la razón práctica propiamente dicha) que opera, una vez resuelto el conflicto, en el intento de optimizar los resultados finales¹⁰⁰. Cuando no aparece una síntesis dentro del conflicto interno de emociones, Spinoza dice que estamos ante la “duda”. En cualquier caso, tampoco somos capaces de elegir esos instintos racionales o racionalizadores: la lógica que se esconde tras la teoría de las preferencias reveladas, y la lógica que hay tras la ley de la utilidad marginal decreciente o de la maximización de resultados son lógicas no-elegidas, implantadas en la evolución de la especie. Se trata de la manera que tiene el cerebro humano de categorizar y ordenar las alternativas¹⁰¹. Mises reconoce que la razón es otro proceso neural en el que no cabe elección:

La razón, el rasgo humano más genuino, es un fenómeno igualmente biológico. No es ni más ni menos natural que cualquier otra circunstancia típica de la especie *Homo sapiens* como, por ejemplo, el caminar erecto o el carecer de pelaje. (Mises (2009[1949]: 211).

⁹⁹ Este vendrá a ser un argumento filosófico central en el posterior pensamiento de Hume y de Kant.

¹⁰⁰ Téngase en cuenta que siguiendo el planteamiento de los instintos pre-programados de Spinoza y de la racionalidad limitada de Simon, tal optimización sólo puede ser de carácter local, o sea, como subraya Samuelson (1993: 145), una configuración sub-óptima del comportamiento. La aspiración teórica a la optimización global denota, parafraseando a Elster (1997: 66-72), una creencia que incurre en el exceso de voluntad y en la hiperracionalidad.

¹⁰¹ Véanse Furlong y Opfer (2009), Robson (2001, y 2002), Greene y Baron (2001), Herrmann-Pillath (1994), Rogers (1994), Cosmides y Tooby (1994a y b), y Frank (1988).

¿Se refieren a lo mismo Hayek y Spinoza cuando hablan sobre la “razón”? Para Hayek (2004[1952]: §5.29) la conciencia y la razón son órdenes espontáneos propios del organismo-individuo, algo así como “mapas” o “planos” (topológicos, no topográficos), y no “planes” de un ente ordenador independiente¹⁰². Quien se guía por la razón, en Spinoza, es quien conoce cómo funcionan los instintos, automatismos, condicionamientos, y procesos de conocimiento y “elección”, es decir, quien conoce esa topología¹⁰³. Spinoza estaría de acuerdo parcialmente con esto dado que, como hemos explicado, para él aquéllos que se guían por la razón son quienes no actúan ciegamente impulsados por las pasiones, sino que conocen sus pasiones y obran con ese reconocimiento de ecuanimidad o conciencia testigo. Sin embargo, Spinoza no llama a esto libre albedrío ni elección tal y como se entiende popularmente. La posible controversia, no obstante, no es tal si tenemos en cuenta que Hayek no se aleja de Spinoza en *El orden sensorial*, donde explica la mente humana como un orden creado por la experiencia (la selección natural y cultural) previa al propio sujeto, esto es, un orden constructor y determinante que es, y produce, una regulación *no elegida* por ningún sujeto-agente¹⁰⁴.

Podemos afirmar que Hayek sí se aproxima a Spinoza cuando habla del orden espontáneo de la mente humana, pero se aleja cuando reconoce que la voluntad libre es una característica del ser humano y que incluso dispone de capacidad para intervenir en la creación de instituciones sociales, esas instituciones que resultan imprescindibles en el sostenimiento de un orden social liberal: así por lo menos la ley, el ordenamiento constitucional y la propiedad¹⁰⁵. Si Hayek dejara todo en manos del orden cósmico que

¹⁰² En *La fatal arrogancia* Hayek critica la extendida y antigua idea de «que dondequiera que se advierta la existencia de un orden debe presumirse la presencia de un ente ordenador» (Hayek 1990[1988]: 59).

¹⁰³ Sin embargo, en su obra de filosofía social Hayek explica con simplificaciones coherentes con la teoría popular (*folk theory*), según la cual guiarse por la razón es elegir libremente, y no condicionados por impulsos y pre-esquemas conativos. Por ejemplo, en uno de los ensayos incluidos en *La contrarrevolución de la ciencia* Hayek parece abrazar las tesis del racionalismo en su exposición del objeto de la praxeología:

Las ciencias sociales en sentido estricto (...) tratan de la acción consciente o reflexiva propia del hombre, de actos de los que puede decirse que una persona realiza en función de una elección entre varias alternativas que se le presentan. (Hayek 2003[1952]: 50).

¹⁰⁴ En realidad, el “sujeto” es una consecuencia del correcto funcionamiento de ese orden espontáneo definitorio de la especie humana. Así, define la mente como

... un orden particular de un conjunto de acontecimientos que tienen lugar en nuestro organismo y que en cierto modo están relacionados con el orden físico de los acontecimientos del entorno (aunque no son idénticos al mismo). (Hayek 2004[1952]: §1.49).

¹⁰⁵ Si no fuera así, todo el pensamiento hayekiano sería irrelevante para una tradición que, aun dentro del *laissez-faire*, no es ajena al constructivismo social, como denota esta reflexión de Milton Friedman:

Spinoza identifica con la Divinidad (la causa incausada, la necesidad pura; en el lenguaje actual el orden espontáneo de la vida humana, la evolución) derribaría el mito de la escuela doctrinal en el que él enmarca su pensamiento: el liberalismo que predica la libertad humana, un mundo visto desde los ojos del ser humano y desde su nicho cognitivo (o sea, un antropocentrismo), una Noosfera, o incluso un Mundo 3 popperiano. Por eso, tanto él como Mises y otros seguidores de la misma corriente doctrinal, se muestran críticos con las visiones “cientifistas” de las ciencias naturales, que intentan explicar al hombre -inviertiendo la hipótesis spinozista- “desde afuera del hombre mismo”, como un objeto más en la naturaleza, y por lo tanto, sometido a las mismas leyes deterministas que el resto de todas las cosas materiales¹⁰⁶. Hayek y Mises no aceptan ese determinismo porque, tal vez, se correría el peligro de negar al hombre su especial “dignidad” dentro de la naturaleza, una dignidad antropocéntrica, mitológica y religiosa.

Así pues, hay dos versiones de la libertad en el pensamiento hayekiano. Una, la de la psicología teórica, que se aproxima al spinozismo y enmarca la libertad dentro de los límites del orden espontáneo de la mente. La segunda, que es la de la teoría política y económica de Hayek, está en el ámbito de la teoría popular de la mente, y presenta la libertad como una característica esencial del único agente racional: el sujeto humano.

2.3.- ¿Qué es el yo? Neurobiología de la conciencia subjetiva

En el *Tratado de la naturaleza humana* Hume dejó escrito que:

... no existe cuestión alguna en la filosofía más abstrusa que la relativa a la identidad y naturaleza del principio unificador que constituye una persona. Hallándonos muy lejos de poder resolver esta cuestión meramente por nuestros sentidos, debemos recurrir a la metafísica más profunda para darle una respuesta satisfactoria. [Hume (1992[1739-1740]: I, iv, 2. p. 127)].

Nuestra sociedad es tal como la hacemos. Podemos modelar nuestras instituciones. Las cosas físicas y humanas limitan las alternativas de que disponemos. Pero nada nos impide, si queremos, edificar una sociedad que se base esencialmente en la cooperación voluntaria para organizar tanto la actividad económica como las demás actividades; una sociedad que preserve y estimule la libertad humana, que mantenga al Estado en su sitio, haciendo que sea nuestro servidor y no dejando que se convierta en nuestro amo (Friedman 1980: 61).

¹⁰⁶ Véase Hayek (1982[1976]) sobre “Los errores de la sociobiología”. Asimismo, sobre los “abusos científicos” del racionalismo constructivista y del positivismo véase igualmente Hayek (2003[1952]).

La metafísica de Hume, a pesar de los esfuerzos, no le llevó mucho más lejos que a asentir:

... cuando penetro más íntimamente en lo que llamo mi propia persona, tropiezo siempre con alguna percepción particular de calor o frío, luz o sombra, amor u odio, pena o placer. No puedo jamás sorprenderme a mí mismo en ningún momento sin percepción alguna, y jamás puedo observar más que percepciones. Cuando mis percepciones se suprimen por algún tiempo, como en el sueño profundo, no me doy cuenta de mí mismo y puede decirse verdaderamente que no existo. [Hume (1992[1739-1740]: I, iv, 6. p. 164)].

El yo es una construcción del cerebro, un epifenómeno real, pero completamente dependiente del funcionamiento adecuado de las complejas interacciones neurales que operan en la supervisión de la homeostasis corporal, el establecimiento de prioridades, y la generación, entre otras cosas, de un modelo de realidad según el cual existe una separación entre el mundo interno y el mundo externo. Este modelo, dentro de un modelo de mayor complejidad, establece las premisas y programas de acción del organismo-en-el-mundo¹⁰⁷. Patricia Churchland (2006: 44) indica que, funcionalmente, el yo «se parece un poco a una aplicación informática de un ordenador, aunque se trata de una que ha evolucionado para crecer y desarrollarse».

La conciencia es el llamado “problema fuerte” de las ciencias que explican la vida y los seres vivos. ¿Por qué de un orden físico de materia y energía ha surgido un orden mental, y de este orden mental ha surgido el pináculo del yo? ¿Por qué existe la conciencia? La conciencia es el ámbito de la subjetividad en la especie humana, y tanto la experiencia consciente como la intuición de un yo que observa, dirige y permanece, son atributos universales de la naturaleza humana. Los humanos no sólo tienen experiencias mentales, sino que además saben que las tienen, e imaginan un yo-observador que es agente unitario y singular:

... parece que es algo que unifica todas mis diversas impresiones sensoriales y todos mis recuerdos (unidad), que asegura estar “a cargo” de mi vida, toma decisiones (libre albedrío) y parece perdurar como una única entidad en el espacio y en el tiempo. (Ramachandran y Blakeslee 1999: 309).

Mosterín (2006: 168), igual que Hume, explica que la conciencia no tiene estabilidad ni permanencia, sino que aparece y desaparece como si algo la activara o desactivara de forma independiente a la voluntad del sujeto. Cuando el organismo está a

¹⁰⁷ Véanse Ramachandran y Hirstein (1997), y Damasio (2005: cap. 5).

punto de dormirse la conciencia se apaga. Cuando está sumamente concentrado en alguna percepción la conciencia se disuelve. El que existan experiencias o pensamientos auto-atribuidos no implica que exista un yo que tenga esas experiencias o que decida tenerlas:

Si por *yo* entendemos la conciencia, no hay uno, sino muchos yoes fugaces y distintos, separados entre sí por períodos inconscientes de no-yo (...) El organismo que soy yo actúa, es un agente, y la aparente coherencia de su actuación parece implicar un principio unificado de decisión, que sería yo. (Mosterín 2006: 188).

De hecho la experimentación neurocientífica reciente sobre procesos mentales asociados a las llamadas experiencias cercanas a la muerte demuestra, en estudios llevados a cabo con ratas de laboratorio, que existe una viva e intensa experiencia mental en animales sin yo, y además sin volición ni motricidad puesto que están clínicamente muertos (Borjigin, *et. al.* 2013). Por esto, y extrapolando a la especie humana, Parfit (1985: 15) sostiene que salvo que *creamos* en alguna forma de animismo, «el sentido común tiene una creencia falsa acerca de la naturaleza de la identidad personal». Tomas Schelling, en una línea de principios teóricos semejante, comenta al respecto:

... es por convención, y no por una convicción científica, por lo que normalmente imputamos continuidad y estabilidad al individuo y a los valores que gobiernan sus preferencias. (Schelling 1985: 362).

Y Hanna Arendt, en *La vida del espíritu*, presenta al yo pensante, encarnado en un organismo humano, como una experiencia discontinua, amarrada a un sujeto-persona corporal y que, en cualquier caso, es una «apariencia entre apariencias» (Arendt 2002[1978]: 287).

Desde el programa de investigación de la psicología evolutiva, el problema fuerte de la vida se explica como cualquier otra faceta físico-orgánica o psicológica, esto es, todos los fenómenos característicos de los universales humanos, desde los automatismos internos de las vísceras, pasando por los movimientos reflejos de los ojos o de las piernas, los impulsos emocionales, hablar o actuar según las normas de una cultura interiorizada, hasta llegar a la conciencia, la identidad personal o lo que

entendemos por comportamiento deliberado e intencional, son estados corpóreos o mentales que tienen una causa biológica. Se trata de adaptaciones filogenéticas para la supervivencia y replicación que se mostraron eficaces cuando surgieron en la evolución orgánica (Pinker 2005). De hecho, los individuos que viven sin yo debido a causas genéticas, enfermedades o lesiones sobrevenidas, muestran una clara incapacidad para sobrevivir de forma autónoma:

... incluso la mera supresión del componente del yo de la conciencia acarrea una disrupción de la gestión de la vida y retorna al ser humano a un estado de dependencia comparable al de un bebé. (Damasio 2005: 197).

La existencia de un yo que se manifiesta a través de la conciencia de la experiencia es un supuesto cuasi-metafísico en las ciencias del comportamiento humano, y también es una experiencia fenomenológica de “algo” según la psicología popular. No obstante, la interpretación dualista del problema fuerte, ya sea en la construcción de la ciencia, en la teoría de la decisión, o en la psicología popular, es un lugar común de la crítica neurocientífica¹⁰⁸. El yo, el “fantasma en la máquina” o “piloto en la nave”, según la neurociencia se trata de un mito culturalmente arraigado que se ha transmitido con éxito mediante la replicación cultural (Ramachandran y Hirstein 1997: 455). Naturalmente, descartada la hipótesis de un “sujeto-que-decide”, el homúnculo cartesiano, queda eliminada la decisión tal y como generalmente se entiende por la tradición racionalista. Esto supone admitir un universo determinista, si bien extraordinariamente complejo, como expone López Corredoira (2008: 263 y 266):

... me parece absurdo pensar que la evolución de las especies ha producido un salto que separa seres libres de no-libres, y que una mutación de un ser no-libre pudo producir un ser libre. Y me parecería absurdo hablar de una libertad por grados (...) O se es libre en una determinada decisión o pensamiento, o no (...) el análisis de la realidad social se relaciona con la fatalidad de la naturaleza humana, parte del “fatum” de un cosmos mecánico cuyas leyes se extienden desde los sistemas más sencillos a las complejidades de la mecánica social. Somos fragmentos de Naturaleza sometidos a sus leyes, y ciertos movimientos mecánicos de la materia en nuestros cerebros, en interacción con su entorno, producen (=“son la causa de”) nuestros pensamientos y decisiones.

¹⁰⁸ Véase el monográfico de *Annals of the New York Academy of Sciences* titulado “The self: From soul to brain”, dirigido por LeDoux, Debiec y Moss (2003). Otro trabajo de neurociencia en este aspecto es el de Metzinger (2003) (puede hallarse una versión resumida en Metzinger 2005).

¿Cómo se explica pues esta ilusión psicológica tan poderosa que caracteriza la psicología humana? Desde el punto de vista de la evolución de la vida es preciso preguntarse por la naturaleza y la función de la conciencia subjetiva en nuestra especie. Richard Leakey (2001: 188-189) precisa en este sentido que:

... la conciencia aparece cuando la simulación cerebral del mundo se hace tan completa que debe incluir un modelo de sí mismo;

del mismo modo,

... un fenómeno mental tan poderoso es probable que haya proporcionado beneficios de supervivencia y que por lo tanto fue producto de la selección natural.

No es posible conocer la vida mental de otros animales, pues el diseño de su cerebro difiere del humano. Sin embargo, sí es posible conocer la experiencia mental de otros humanos, e incluso la de mamíferos con un cerebro parecido al humano, como el caso del chimpancé¹⁰⁹. Esas capacidades representacionales de las experiencias de otras mentes pierden significación conforme aumenta la distancia homológica entre las especies, así es en el caso del chimpancé, homológicamente cercano, frente al murciélago, mamífero también, pero homológicamente más alejado¹¹⁰. Sin embargo, sobre la base del conocimiento derivado de la experimentación neurológica con otros animales, y sobre la psicología del desarrollo y la investigación de enfermedades neurológicas, se pueden plantear algunos resultados concluyentes sobre la naturaleza y la funcionalidad evolutiva de ese fenómeno mental al que llamamos yo. Pasamos a examinarlos.

¹⁰⁹ Esta operación neurológica consiste en la transcripción-decodificación de la información recibida cuando se percibe a otro individuo (hipótesis de la simulación a través de neuronas espejo), así como a través de su evaluación lógica y contextual (hipótesis de la inferencia a través de una teoría de la mente) (Gallese y Goldman 1998). En las dos últimas décadas se ha comenzado a confeccionar el mapa de circuitos neurofisiológicos y áreas cerebrales implicadas en la realización de estas operaciones, fundamentalmente para el caso humano y de algunos otros primates. Véase al respecto Frith y Frith (2003), Oberman y Ramachandran (2007), y Rizzolatti y Sinigaglia (2006).

¹¹⁰ No obstante, aunque desconocemos los *qualia* (experiencias en bruto que afloran a la conciencia del murciélago; por ejemplo, las representaciones derivadas de la ecolocalización), no debe existir discrepancia en lo esencial cuando los comportamientos ante estímulos y situaciones diversas son similares y predecibles con éxito desde la perspectiva de un observador humano (Dennett 1971).

2.3.1.- El yo es un proceso neural de alto nivel

El atributo de la identidad personal y del conocimiento auto-reflexivo es un rasgo exclusivo de la especie humana. Sin embargo, dado que los cerebros de los mamíferos, y especialmente de los primates, tienen semejanzas anatómicas y fisiológicas, es posible que la especificidad de las experiencias mentales humanas sean más una cuestión de grado (más cantidad de neuronas y conexiones más complejas) que de naturaleza y estructura (Churchland 2006: 44). La psicología comparada ha mostrado una elevada semejanza en la expresión de emociones y comportamientos entre la especie humana y otras especies de mamíferos¹¹¹. De hecho, a los humanos pueden llegar a entablar una proto-comunicación con otros animales de diferentes especies, y son capaces de intuir los estados psíquicos tanto de animales domésticos como de otros animales salvajes a través de la comunicación no verbal¹¹². ¿Qué tipo de cerebro se necesita para que surja un yo?

Las neurociencias consideran que el yo carece de realidad óptica y epistémica, y que se trata de un proceso cognitivo de alto nivel que integra, como intuyera Hume, toda una variedad de procesos y subprocesos mentales en un mismo interfaz¹¹³. Así, existe un proceso neural que registra unificadamente el acontecer neurológico de otros procesos (no todos con el mismo nivel de conciencia, dado que existen operaciones conscientes e inconscientes) de tipo homeostático, motivacional, memorístico, emocional y ejecutivo. Este registro de información se lleva a cabo mediante representaciones generadas en un modelo del organismo integrado en otro modelo que representa a los organismos y objetos “externos”, es decir, al “mundo”. Estos modelos han de ser coherentes y estar sincronizados, y por eso, en determinadas circunstancias, el cerebro genera ilusiones, confabulaciones y reconstrucciones sesgadas de la realidad,

¹¹¹ El trabajo clásico a citar en este aspecto es Darwin (1998b[1872]). Dos trabajos contemporáneos en la línea teórica del darwinismo son los de Daly y Wilson (2005 y 1999).

¹¹² En *El anillo del rey Salomón* Konrad Lorenz (2002[1949]) defiende la tesis de un proto-lenguaje animal con una base biológica idéntica a la humana.

¹¹³ El cerebro humano funciona mediante procesos jerárquico-funcionales anidados, igual que una organización bien sincronizada, acoplada y sin desajustes (véase Sanfey, *et. al.* 2006: 109). El yo es una manifestación psicológica de ese orden, carente de una localización física específica:

... sabemos que son diversas partes del cerebro las que intervienen en la creación del yo, sin embargo, no hay un emplazamiento material específico del *self* o del “yo” en el cerebro (...) El cerebro crea la unidad del yo mediante la producción de una jerarquía anidada de significado y propósito, en la que los niveles del yo, y las múltiples partes del cerebro que contribuyen a producir el yo, están anidadas unas en los otros niveles de jerarquía (...) nos experimentamos a nosotros mismos como algo unificado porque nuestros significados y nuestras acciones están unidas dentro de ese yo anidado. (Feinberg 2001: 149).

a fin de forzar y reforzar esa coherencia y sincronía. Lo hace interpretando y dando sentido a la acción que surge de los procesos internos, así como rellenando huecos de información cuya ausencia pondría en peligro la integración de todo el entramado. Cuando todo funciona bien el organismo está capacitado para sobrevivir y reproducirse. Cuando hay algún problema de acoplamiento entre ambos modelos, como ocurre con la enfermedad de la anosognosia, el cerebro está enfermo o dañado, entonces el organismo se muestra incapaz de sobrevivir y reproducirse autónomamente, o bien sobrevive con graves carencias que repercuten en su eficacia reproductiva¹¹⁴.

¿Cómo funciona este diseño neurológico capaz de producir un yo? Una forma de contestar la pregunta es comparar un cerebro en situación de no-yo con un cerebro cuyo estado de conciencia normal es el de una mente-sujeto. Para ello puede ser esclarecedor comparar el funcionamiento de un cerebro que está en estado de meditación profunda con otro que está en vigilia y alerta reflexiva. En esencia, y muy sintéticamente, un cerebro produce la sensación de un yo-agente a través de los siguientes sistemas neurológicos (D'Aquili y Newberg 1999: 34 y ss.; Ramachandran y Blakeslee 1999: 308 y ss.)¹¹⁵:

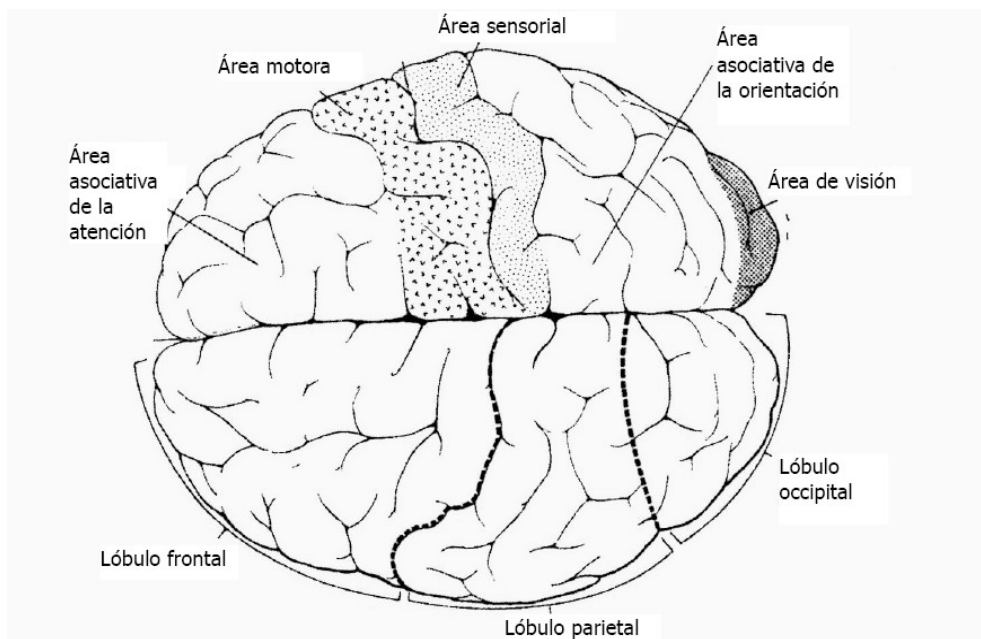
- a) Área asociativa de la atención: ubicada en el córtex prefrontal, es la responsable de integrar todas las sensaciones y percepciones intero y esteroceptivas del organismo. Toda la información sobre los estados internos del cuerpo y sobre lo que ocurre afuera del cuerpo y que pueda tener alguna repercusión en él se centraliza en esta área, desde donde se distribuye la información para ser coordinada y organizada de acuerdo a prioridades homeostáticas.

¹¹⁴ La falta de acoplamiento de los modelos interno y externo de representación y semántica de los acontecimientos se aprecia muy bien en los problemas neurológicos relacionados con la empatía, y los déficits de cognición social en especies de primates sociales. En un trabajo experimental realizado con monos en libertad, Brothers (1989) concluye que los animales con daños en el circuito prefrontal-amígdala son capaces de realizar tareas complejas, como la recolección de alimentos y otras necesarias para su manutención y supervivencia; sin embargo, se muestran incapaces de relacionarse y coordinarse con monos de su grupo, al no poder comprender las emociones que éstos expresan. Los monos discapacitados incluso huyen ante la presencia de compañeros de grupo que les manifiestan gestos que el resto de individuos sin daños neurológicos entienden perfectamente como señales amistosas.

¹¹⁵ Véanse Figuras 2.1. y 2.2. sobre la organización funcional del cerebro humano, así como la Figura 2.3. para una presentación elemental de su anatomía .

- b) Sistema límbico: integrado por las amígdalas y sub-zonas de los lóbulos temporales. Es el responsable de establecer los impulsos automáticos conducentes a la activación de procesos emocionales que, en la esfera consciente, se transforman en sentimientos conducentes a la motivación y a la acción en orden a preservar la estabilidad homeostática del organismo (Damasio 2005). Está conectado con el área asociativa de la atención en un flujo bidireccional de información, así como con el área motora (ubicada en una sub-zona del lóbulo frontal), cuando se activan procesos autónomos de acción que no afloran a la consciencia, como los que p.ej. producen el estornudo o la acción de escapar ante la súbita presencia de una amenaza.

Figura 2.1.- Vista superior del cerebro humano

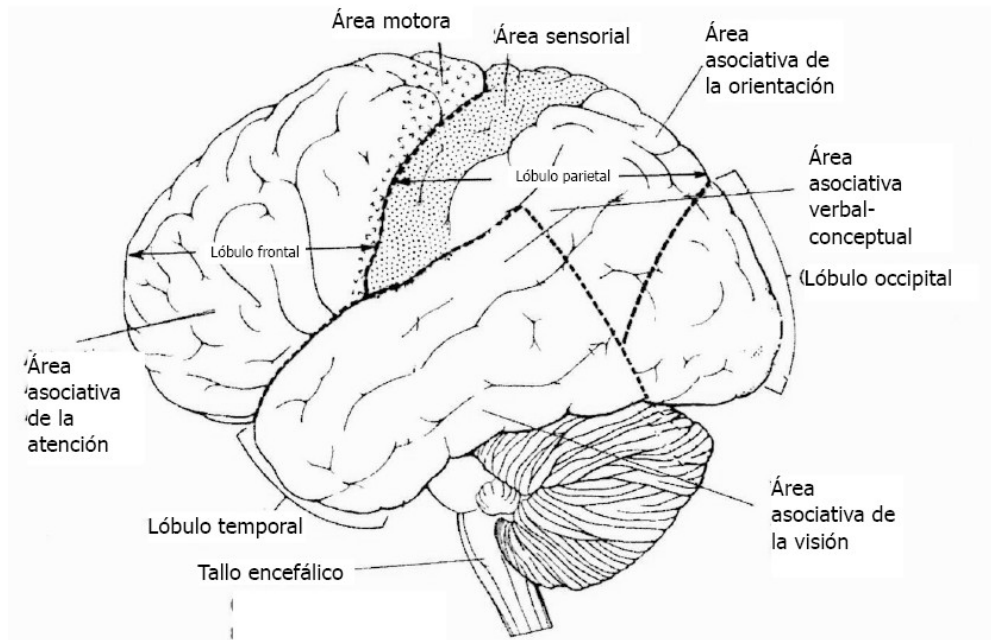


Fuente: D'Aquili y Newberg (1999: 29)

- c) Área asociativa de la orientación: ubicada en una sub-zona del lóbulo parietal superior. Responsable de procesar la información acerca del tiempo y del espacio, o sea, de ubicar y orientar al organismo en un mundo dentro del cual existe él, creando la sensación de qué es y dónde

está lo propio, y diferenciarlo de lo que no es propio. Sin este proceso, el organismo no halla diferenciación entre el yo y el resto del mundo.

Figura 2.2.- Vista lateral del cerebro humano



Fuente: D'Aquili y Newberg (1999: 29)

- d) Área de la memoria: está ubicada en el hipocampo y conectada con las amígdalas. Aquí se almacenan los recuerdos en el sentido de vivencias recordadas en primera persona, y no información “sin sujeto”. Por ello la información almacenada está acompañada de un “aroma emocional” que la hace propia, y por lo tanto vivida en primera persona y autobiográfica. Esta biografía ha de ser coherente, pues de lo contrario el individuo carecería de una identidad unificada. Es por esto que la memoria rellena huecos de información y reconstruye “interesadamente” el pasado (Sacks 2005; Eakin 2001; Schacter 1995).
- e) Área de asociación verbal-conceptual: ubicada en el lóbulo parietal inferior y tiene múltiples conexiones con el centro del lenguaje (emplazado en un amplia zona distribuida entre el lóbulo temporal

izquierdo -el área de Broca- y partes de los lóbulos frontales y temporales). Contiene las conexiones responsables de la generación de representaciones conceptuales abstractas y categorías cognitivas, y de asociarlas a palabras del lenguaje. Esta función permite “reificar” experiencias mediante palabras, y entre otras la experiencia de ser un sujeto, o yo¹¹⁶. Está comprobado que la principal diferencia entre el cerebro humano y el de otros mamíferos no humanos estriba en que en ellos este área se halla mucho menos desarrollada.

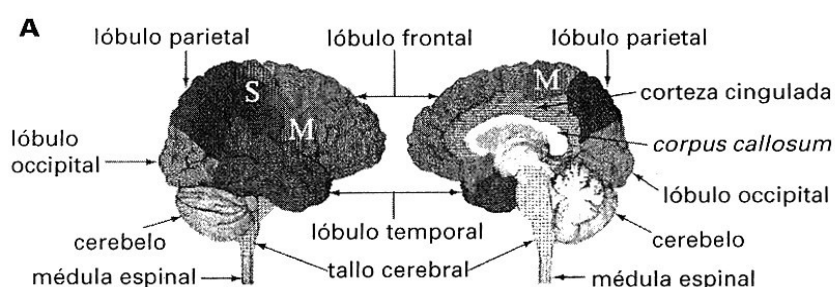
- f) Área ejecutiva: ubicada en las zonas motoras del giro cingulado anterior y zonas suplementarias de los lóbulos frontales. Se trata de un área moduladora que genera actividad en el organismo y evita la respuesta motora desordenada ante cualquier estímulo, así como la de ser un yo pasivo e incapaz de intervenir. La sensación de agencia, libre albedrío y voluntad, sobre la base de la reflexión y de la deliberación conscientes, es en realidad una «disposición a actuar condicionada» (Ramachandran y Blakeslee 1999: 312), que incluiría una capacidad de interrumpir procesos ya iniciados en escalas de tiempo muy pequeñas, que contribuyen a generar la impresión de que se lleva el control. En conexión con el área asociativa de orientación, las amígdalas y el área motora, el cerebro genera un modelo representativo del mundo externo, así como otro modelo, integrado en el primero, en el cual se representa el yo-organismo. Para que se pueda “tomar decisiones” el cerebro valora automáticamente objetivos, medios, posibilidades lógicas de realizar acciones (p.ej. la de parar un tren en movimiento con las manos, o de volar hasta la luna), y probabilidades de que ocurran acontecimientos (que esta tarde llueva, o que mañana bajen las cotizaciones bursátiles). Con toda esa información, integrada en los modelos del mundo y del yo-

¹¹⁶ Uno de los contenidos de las experiencias mentales es el de la experiencia de unicidad y permanencia individual. Se trata de una “reificación”. Con la palabra “reificar” se alude al tratamiento de las representaciones abstractas que son los *qualia* otorgándoles la categoría de ser algo real. En este sentido, podemos decir que *cosificar* vale como sinónimo de *reificar*. La consciencia, que incluye la cosificación de un sí mismo, supondría que se tiende a considerar “real” la representación mental del cuerpo físico, y además la de un sujeto abstracto que representa a su vez la representación mental del cuerpo.

mismo, integrados además ambos modelos entre ellos, y recibida información del resto de subsistemas (especialmente del sistema límbico, que marca la orientación de las acciones a través de la motivación), el área ejecutiva genera órdenes motoras para el organismo. Estas órdenes motoras no son equiparables a un modelo de racionalidad de alta potencia, como la metáfora del demonio de Laplace, sino heurísticas muchas veces pre-programadas (genéticas o aprendidas), frugales y operativas, que resultan eficaces sin aspiraciones ni necesidad de optimización (Gigerenzer 2000)¹¹⁷.

Asumiendo la experiencia de ser un sujeto como un interfaz de experiencia unificada, entidad diferenciada y autónoma que perdura en el espacio y en el tiempo, una biografía coherente, dotado todo ello de capacidad libre y autónoma de acción (agencia), la interacción de todos estos dispositivos neurales es la que hace posible la producción de esa ilusión. Cuando los módulos y circuitos no están adecuadamente construidos, se encuentran desactivados, desconectados, o sufren patologías (como las derivadas de enfermedades, tumores y cirugías), el cerebro no es capaz de generar estas experiencias cuya integración es lo que llamamos el yo. Para entender por qué la evolución ha dado lugar a la ilusión del yo se puede hacer el experimento mental de una ilusión de no-yo, o saber lo que sucede en casos de no-yo o yoes patológicos.

Figura 2.3.- Anatomía del cerebro humano



¹¹⁷ En el *Essai philosophique sur les probabilités* (1814) Laplace planteó la posibilidad de un ser hiperracional capaz de concebir todas las posibilidades del universo en un solo momento, y actuar deterministamente en la consecución de sus objetivos. La alegoría de Laplace es un experimento mental sobre un cerebro dotado con capacidades cognitivas y computacionales ilimitadas. Obviamente, no es el caso del cerebro *Homo sapiens*.

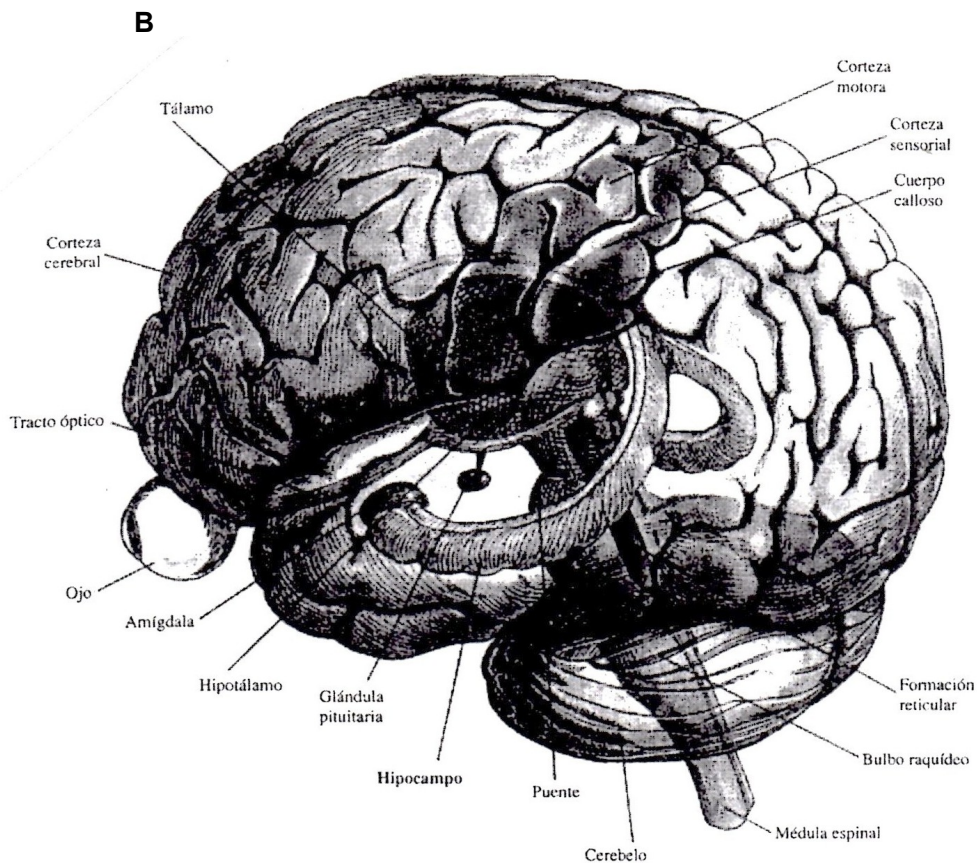


Figura 2.3.A: Esquema lateral y sección central del cerebro. Fuente: Damasio (2005: 273).

Figura 2.3.B: Esquema lateral detallado. Fuente: Ramachandran y Blakeslee (1999: 40).

2.3.2.- El funcionamiento del cerebro en situaciones de pseudo-yo, no-yo y yoes patológicos

El cerebro del neonato humano se caracteriza por estar en proceso de rápida construcción y organización a través de programaciones epigenéticas interaccionando con estímulos ambientales. Al momento de nacer su cerebro tiene apenas la cuarta parte del tamaño que acabará teniendo cuando llegue a ser adulto (Gould 2010[1977]: 79). Una de sus propiedades es que el cerebro aún no suficientemente formado del niño pequeño no dispone de la arquitectura neural necesaria para producir esos modelos del yo y del mundo con que los adultos atribuimos acciones a un sujeto rector en un mundo que tiene un orden coherente y funcional, y que está habitado por otros sujetos semejantes al yo. Durante los dos primeros años de vida se van a desplegar las programaciones epigenéticas que incrementarán de forma extraordinaria tanto el

volúmen como la complejidad del cerebro infantil¹¹⁸. Estos procesos espontáneos acabarán posibilitando la aparición del yo, y paralelamente a ese epifenómeno aparecerán a la conciencia otros yoes diferenciados cuya vida subjetiva no será del todo desconocida, pues a partir de los tres años se va a ir desarrollando la capacidad neuropsicológica de “leer” los estados mentales de otros individuos, inferir sus pensamientos, intenciones y expectativas, y comprender sus sensaciones, emociones y estados de ánimo.

El tiempo durante el cual se conforma la autoconciencia se caracteriza por la psicología de un no-yo o pseudo-yo que Piaget denominó “fase egocéntrica” en *El nacimiento de los sentimientos morales en la infancia* (1932). El cerebro infantil, aún no suficientemente desarrollado, no es capaz de discernir la individualidad, y produce una experiencia solipsística de no-yo o pseudo-yo en la que no se conciben las categorías de lo “interior” y “exterior” del organismo. Asimismo, a esas edades en las que el yo no aparece de manera clara y distinta, como conciencia reflexiva e independiente relacionada con el mundo externo, los niños y niñas no pueden inferir ni razonar acerca de los estados mentales (inobservables) de los demás. No han desarrollado una teoría de la mente porque los dispositivos neurales anatómicos y funcionales que la posibilitarán todavía no existen¹¹⁹. Hasta que eso no empiece a ocurrir, ya después de los 3 años, no son capaces de superar el test de la falsa creencia, e interpretan todo lo que perciben, desde objetos, animales y plantas, hasta otros niños, exclusivamente desde su propia y única perspectiva:

... es incapaz de diferenciarse del resto del mundo y, en ese sentido, todavía no se ha separado de los demás ni de los objetos (...) siente que los otros comparten su dolor o su placer, que inevitablemente deben comprender las palabras que apenas masculla, que su perspectiva es compartida por todas las personas y que hasta los animales y las plantas participan de su conciencia. Así, cuando juega al escondite cree que, si no ve a los demás, ellos tampoco podrán verle, porque su egocentrismo le impide reconocer que el punto de vista de los demás es diferente del suyo. (Howard Gardner, citado en Wilber 2000: 38).

¹¹⁸ El 90% de las sinapsis de la corteza cerebral se forman en el período que transcurre entre el nacimiento y los dos primeros años de edad (J. P. Changeux, citado en Levi Montalcini 2000: 147).

¹¹⁹ Los experimentos con técnicas de neuroimagen llevados a cabo por Rebecca Saxe y sus colaboradores, están permitiendo conocer la interrelación existente entre la psicología del desarrollo infantil y las transformaciones, neuroanatómicas y neurofuncionales, relacionadas con la aparición de la intersubjetividad en la infancia. Específicamente, el equipo de Saxe está especializado en la neuropsicología de la teoría de la mente, que permite atribuir estados mentales a otros yoes de manera no solipsística. Véase Gweon, *et. al.* (2012).

Esta perspectiva egocentrada, que no es patológica, aún no capacita al individuo infantil para una supervivencia autónoma en ningún aspecto relevante, y específicamente no posibilita relacionarse en una vida de interacción social compleja (Warneken y Tomasello 2006).

Otra modalidad de cerebro sin yo, o con un pseudo-yo no necesariamente patológico, es la de aquéllos cerebros predispuestos congénitamente a ciertas experiencias místicas de ensimismamiento, o entrenados para la meditación contemplativa. Los estados místicos se caracterizan por una experiencia de no espacio, no tiempo, no objetos ni sujetos, no sensaciones ni pensamientos, no separación entre interior y exterior, ausencia de intenciones, intereses o deseos. La vivencia es la de un observador metafísico contemplando la vacuidad, y más allá tampoco eso, sino solamente una oceánica sensación de ser unitario y absoluto, sin espacio ni tiempo (o en espacio y tiempo infinitos), una “pura conciencia” (D'Aquili y Newberg 1999: 183).

En los estudios de neuroimagen realizados por D'Aquili y Newberg (1999) con monjes budistas y monjas franciscanas, tanto en estados normales de vigilia como en estados de meditación profunda, se aprecia durante los estados místicos una desconexión de la actividad neural (una “deafereenciación”) en las áreas relacionadas con la experiencia del yo¹²⁰. Esta deafereenciación afecta especialmente a las zonas posteriores, derecha e izquierda, del lóbulo parietal superior, y también del lóbulo parietal inferior, zonas en las que se localizan las áreas asociativas verbal-conceptual y de orientación. Si se bloquea la actividad neural en esta zona, se inhibe la experiencia de tiempo y de espacio, se disuelve la separación entre el cuerpo y el resto del mundo, y consecuentemente se experimenta un yo atemporal infinitamente abarcador. El cerebro es incapaz de reconocer dónde termina lo propio y dónde empieza lo que no es propio. La conclusión a la que llegan estos investigadores es que

... incluso desde la perspectiva de la prioridad de una realidad material externa, parece que el yo, o ego consciente, no tiene un estatus *a priori*, sino que se trata de un constructo práctico surgido de la evolución física. (D'Aquili y Newberg 1999: 189).

¹²⁰ Otros estudios neurocientíficos sobre el funcionamiento del cerebro y la fenomenología asociada a los estados místicos son los de Beauregard y Vincent (2006) y Cahn, *et. al.* (2010).

El yo es una reificación generada en el área de asociación verbal-conceptual, de manera similar a la asimilación de cualquier otro concepto socialmente construido, como los de azul, libertad o felicidad. De modo que

Esta reificación de las diversas funciones mentales/cerebrales conscientes es el yo (...) el yo consciente es un producto de procesos evolutivos, que está siempre al tanto de “lo otro”, y que es incuestionablemente diferente de la conciencia pura (no-dual). (D'Aquili y Newberg 1999: 189; paréntesis añadido).

Una pregunta importante sobre los estados espontáneos o inducidos de no-yo es la de qué función evolutiva puede tener un cerebro capacitado para deaferenciarse. Muchas prácticas litúrgicas y contemplativas de la mayoría de las tradiciones religiosas, como es el caso del budismo zen, tienen como objetivo reducir el sentido subjetivo del yo. Curiosamente, como explica Kolm (1987: 244), estos rituales y prácticas que estimulan el estado psicológico de no-yo, son eudemonistas en el sentido de que mejoran el estado de ánimo y la alegría vital de los practicantes. La mentalidad occidental, que considera la desidentificación del ego como un cese del egoísmo y un refuerzo del altruismo, no parece considerar, como sí ocurre en las tradiciones místicas, especialmente orientales, que el no-yo, que es la experiencia de ecuanimidad por antonomasia, produce un refuerzo de la disposición al comportamiento altruista y una reducción de la angustia existencial (causada por la incertidumbre egocéntrica ante lo desconocido y la impermanencia, entre otras posibilidades)¹²¹.

Se ha constatado que hay individuos predispuestos genéticamente a tener este tipo de experiencias místicas (que en ocasiones pueden resultar patológicas), debido a su propensión a sufrir epilepsias localizadas en los lóbulos temporales, y/o disfunciones en la interacción entre el sistema límbico y dichos lóbulos (Ramachandran y Blakeslee 1999: cap. 9). Pero la capacidad general de acceder a estados deaferenciados, ya sea espontánea o resultado de prácticas culturales, requiere la capacidad psicológica previa para transformar la experiencia normal del yo-relacionado-con-el-mundo en una experiencia de yo-místico. Esta disposición neuropsicológica puede resultar de una

¹²¹ En las tradiciones místicas contemplativas lo opuesto al yo no es el tú, sino el no-yo. La compasión universal y la atenuación de la neurosis se consigue con la deaferenciación a través de un hedonismo no consistente en la búsqueda de nuevos placeres, sino con la evitación del dolor mediante la *aponía* (ausencia de dolor corporal) y la *ataraxia* (ausencia de sufrimiento psíquico), dos de las metas del particular eudemonismo epicúreo originario.

disfunción filogenética, una misteriosa readaptación de otras funciones mentales, o también de ventajas adaptativas asociadas a los efectos que pueda tener el estado de no-yo en la reducción del estrés, la concentración para la resolución de problemas, la mejora de la salud física y mental, la actitud de apertura relacional ante los demás y la cohesión grupal en torno a sistemas socio-culturales de tipo religioso¹²².

En cualquier caso, el yo cumple funciones adaptativas en la especie humana (incluso admitiendo las probables consecuencias benéficas de la deaferenciación ocasional y controlada), y una buena prueba de ello es que un organismo humano en estado permanente de no-yo viene a ser como un cerebro desconectado, incapaz de mantenerse vivo en un organismo vivo, algo así como un animal carente de emociones, o profundamente dormido en medio de un mundo peligroso y sin protección. La muerte por inanición o por ataque de depredadores fue el destino inmediato más probable del portador de un cerebro deaferenciado en el Pleistoceno, salvo que otros, los que sí experimentaron la vida desde el yo-relacionado-con-el-mundo, lo alimentasen y protejieran.

Algo parecido puede decirse de aquellas enfermedades neurológicas que producen un yo disfuncional y patológico. Existen numerosas enfermedades neurológicas que repercuten en la incapacidad de tomar decisiones o que llevan al organismo a actuar “sin control”. Tales enfermedades guardan relación con daños o disfunciones en al menos uno de los módulos y/o conexiones modulares involucrados en la generación de la identidad personal. Los pacientes con daños o epilepsias en los

¹²² Independientemente de las ventajas asociadas al grupalismo religioso en cuanto al refuerzo de los procesos de selección en varios niveles (Wilson 2002), está constatado experimentalmente que la práctica habitual de la meditación explica ciertas transformaciones psicológicas y fisiológicas en el individuo practicante. Entre las primeras están la mayor disposición al comportamiento pacífico y tolerante, a los sentimientos de compasión y amor al prójimo, mayor capacidad de atención y creatividad, y resiliencia ante las dificultades. Asimismo, se han demostrado otros cambios de tipo somático como son la disminución del ritmo cardíaco y de la presión sanguínea, la reducción de la frecuencia respiratoria, aumento de la resistencia galvánica de la piel, mayor oxigenación de las células, reducción del lactato en el riego sanguíneo, incremento de la actividad autoinmune, y otras alteraciones endocrinas a largo plazo. En cuanto disposición emotiva, se ha comprobado una mayor tendencia en los practicantes a experimentar emociones positivas, superior sosiego mental, ecuanimidad y desapego, y en general, reducción de la ansiedad y un estado vital de mayor bienestar y autorrealización. Parece verosímil por lo tanto postular ciertas ventajas adaptativas relacionadas con la deaferenciación. Véanse al respecto los trabajos de Chiesa y Serretti (2010), Rosenkranz, *et. al.* (2003), Lutz, *et. al.* (2009), Brefczynski-Lewis, *et. al.* (2007), Davidson (2003), Martín Asuero, *et. al.* (2005), Urry, *et. al.* (2004). Una presentación de la literatura principal puede encontrarse en García Campayo (2008).

lóbulos temporales (especialmente el caso del lóbulo izquierdo), amígdalas, el córtex inferotemporal, área de Wernicke y otras estructuras asociadas con el sistema límbico (como el giro cingulado) experimentan graves desórdenes en sus comportamientos, así como en el control de las acciones y en la experiencia de ser individuos autoconscientes con una identidad personal¹²³. Del mismo modo, algunos de ellos experimentan alucinaciones muy intensas (incluidas experiencias extracorporales, o sensación de omnipotencia y omnisciencia absolutas) que les impulsan a actuar de una manera frecuentemente temeraria, carente de coherencia o autocontrol, al margen de las normas sociales o morales, y disociada de su personalidad y creencias interiorizadas antes de producirse la patología. Estos son los casos de enfermedades como las siguientes¹²⁴:

- a) La esquizofrenia, en la que el individuo siente cómo fuerzas y voluntades ajenas a él dirigen su comportamiento;
- b) la anosognosia o la somatoparafenía, en las que el paciente no identifica partes de su cuerpo como propias (suelen afirmar que pertenecen a otras personas), y por lo tanto no siente responsabilidad de las acciones que estas áreas de su cuerpo realizan;
- c) el mutismo acinético, que es una enfermedad en la que el individuo, plenamente consciente de su entorno, sin embargo permanece constantemente inmóvil, como sin voluntad o capacidad de hacer nada en absoluto;

¹²³ Churchland (1996: 186-187) explica que cuando se analiza experimentalmente el comportamiento de personas con patologías neurológicas respecto a individuos considerados sanos, lo que se aprecia comparativamente es que con frecuencia el comportamiento de los primeros está “fuera de control”. En el caso de los individuos sanos, su comportamiento está mejor orientado a los intereses de supervivencia y replicación tanto a corto como a largo plazo, toman decisiones y planifican de manera más prudente, informada y razonada. Existen diversas estructuras neurales directamente implicadas en el funcionamiento normal de un cerebro sano, siendo las áreas cerebrales críticas en la situación patológica de un organismo (un yo patológico irreversible, en el caso humano) el córtex frontal ventromedial y la amígdala.

¹²⁴ Para una información más detallada acerca de todas estas enfermedades y otras más relacionadas con fallos en la conciencia de la voluntad y la toma de decisiones véase Ramachandran y Blakeslee (1999).

- d) el síndrome de Cotard, en el que el paciente se siente tan desligado emocionalmente del mundo y de sí mismo, que puede llegar a afirmar tajantemente que está muerto; o
- e) el “síndrome de la mano extraña”, en el que una o las dos manos del paciente realizan movimientos y operaciones incluso complejas, pero de forma completamente desligada de sus pensamientos e intenciones conscientes (algunos pacientes incluso afirman que su mano tiene mente propia, o que es movida por alguien diferente y externo a ellos).

A estas patologías habría que añadir las enfermedades que deterioran o eliminan los recuerdos autobiográficos y las adscripciones a grupos sociales, como las demencias y la enfermedad de Alzheimer (Jetten, *et. al.* 2010). En estos casos de pérdida irreversible de la memoria, el individuo olvida progresivamente las funciones motoras, los recuerdos y sus adscripciones psico-sociales. El individuo deja de tener identidad personal y social, lo que le expone a la completa dependencia de otras personas, además de ser un factor altamente estresante para el organismo que a la larga acaba dañando otros órganos y funciones vitales (Jetten, *et. al.*, 2010: 409). El autismo no es una enfermedad que podamos considerar generadora de un yo patológico, sin embargo, algunas características de la enfermedad, como el ensimismamiento, la incapacidad de comprender la subjetividad ajena, o la falta de coherencia centralizada y la lentitud en el proceso de integración de la información sensorial, contribuyen a la expresión de un yo con graves problemas de integración social, “desconectado” de las relaciones sociales, y vulnerable a estrategias de engaño y explotación ajenas (Oberman y Ramachandran 2007). Esto se debe a que los enfermos de autismo carecen de los dispositivos neuropsicológicos que capacitan para una teoría de la mente, o tales dispositivos no están adecuadamente desarrollados (Baron-Cohen 1995; Baron-Cohen, *et. al.* 2000).

2.4.- Crítica al concepto estándar de racionalidad en las ciencias sociales

Hayek define la racionalidad como una facultad humana que apareció cuando las representaciones del mundo surgidas de las reglas (genéticas y culturales) con que la mente clasifica y ordena las percepciones, pasaron a formar parte de las variables de un modelo de cognición que permite predecir y anticipar el futuro (Hayek (1982[1976]: 268-269). Es decir, para Hayek la racionalidad consiste en un proceso marco de ordenación y representación de la información producto del cual se llevan a cabo acciones basadas en inferencias y planes sobre el futuro. La acción humana es fruto de un orden y se engarza en otros órdenes de mayor complejidad. Se trata de sistemas neurológicos surgidos de la evolución, y que estructuran y ordenan la información tanto de los procesos orgánicos individuales como del mundo en el que los individuos viven e interaccionan. Por ello la conducta humana no es contradictoria ni aleatoria, sino que tiene coherencia y un sentido comprensible. En la teoría hayekiana de la mente los diferentes subsistemas neurales se integran en un modelo que permite seleccionar comportamientos en función de expectativas sobre los objetos del mundo exterior y sobre otros individuos. Dentro del racionalismo filosófico estricto, dicha competencia integradora es la facultad del raciocinio y, dependiendo de los modelos teóricos, la exigencia del raciocinio será de mayor o menor rigurosidad, y afectará a más o menos elementos integrantes de la toma de decisiones.

La tradición racionalista en filosofía teoriza la acción humana como una operación normativa consistente en actuar sobre la base de una decisión e intención previas a la acción (Hardy-Vallée 2007: 939). En este sentido, Max Weber (1969[1922]: 5) precisa que la acción humana se define por estar “enlazada” a un sentido subjetivo. Según la teoría racionalista de la decisión toda acción humana es intencional y dotada de significado subjetivo. La acción será además racional si el agente está motivado por razones, de modo que las creencias y deseos sean ingredientes constitutivos de la racionalidad, que dan sentido y justifican la acción racionalmente. Se precisa pues una evaluación acerca de la coherencia lógica entre intenciones, acciones y planes (Hardy-Vallée, 2007: 947). De este modo, las razones que motivan la acción de un animal

racional son actitudes proposicionales (o sea, motivaciones que pueden argumentarse lingüísticamente), y no impulsos inconscientes o tropismos.

En un sentido débil del concepto, la racionalidad de la acción viene apoyada en deseos y en creencias, siendo las creencias lo más próximo a la verdad lógica objetiva. Esto no excluye que puedan ser impremeditamente erróneas, o resulten contradictorias, incoherentes, absurdas o indemostrables (Davidson 1982: 318). En un sentido fuerte del término, tanto los deseos como las creencias, y las reglas con que se integran deseos y creencias en orden a producir una decisión, han de ser verdades lógicas demostrables y coherentes¹²⁵. Así, un conjunto coherente de preferencias, unas creencias bien formuladas en términos bayesianos, y un procedimiento lógico y bien calculado de realización de un plan, son los resultados de una mente racional, y no sus causas o precondiciones (Searle 2000: 50). Además, y dentro de una acepción más estricta aún, ni creencias ni deseos son pre-condiciones causales suficientemente explicativas para que una acción se considere racional, dado que se requiere también la racionalidad de los fines. Actuar sólo motivado por deseos y sobre la base de creencias, sin tener en cuenta los condicionantes normativos señalados consiste, más bien, en

... casos extravagantes y típicamente irracionales. Son casos en los que, por ejemplo, el agente ha caído en las garras de una obsesión o una adicción. (Searle 2000: 27).

Esta orientación teórica de la racionalidad, tanto en su sentido fuerte como débil, está influida por el racionalismo cartesiano, según el cual sólo es racional la dimensión consciente, lógica y proposicional de la mente, y dado que sólo es consciente, lógico y proposicional el pensamiento, por ello sólo es racional aquello expresable y justificable mediante pensamientos (Davidson 1982: 323). Siendo tales las restricciones, sólo puede ser racional el comportamiento consciente e intencional del ser humano: ni los comportamientos inconscientes humanos son racionales ni, por supuesto, ningún comportamiento animal lo es. Obviamente, dado que sólo el ser humano es capaz de pensar y expresar lingüísticamente juicios y evaluaciones, actitudes y motivos, según esta orientación sólo los humanos son animales racionales.

¹²⁵ Una digresión mucho más detallada sobre la acepción fuerte de racionalidad y sus implicaciones metodológicas se halla en Colman (2003: 140 y ss.)

Según Weber, este modelo ideal de conducta racional con arreglo a fines expone cómo se desarrollaría la acción humana si se llevase a cabo con todo el rigor lógico posible. Sin embargo, matiza que la completa racionalidad en la acción con arreglo a fines es un caso límite, un modelo ideal para su manejo teórico:

La acción *real* sucede en la mayor parte de los casos con oscura semiconsciencia o plena inconsciencia de su “sentido mentado”. El agente más bien “siente” de un modo indeterminado que “sabe” o tiene clara idea; actúa en la mayor parte de los casos por instinto o costumbre. Sólo ocasionalmente (...) se eleva a conciencia un sentido (sea racional o irracional) de la acción. (Weber 1969[1922]:18).

Otra acepción de la racionalidad, en este caso apoyada en la lógica matemática, presenta la razón como cálculo de costes y beneficios individuales. Según este enfoque el cálculo racional es el atributo fundamental de la naturaleza humana¹²⁶. La actividad mental racional se considera *causa* del tipo de comportamiento característico del ser humano. De este modo, la metáfora del *homo oeconomicus* consiste en un modelo individualista de racionalidad práctica, delimitado en los estados psicológicos del individuo como foco autónomo e independiente de experiencias (Boland 1982).

En la visión economicista de la racionalidad práctica el comportamiento individual es consistente con el propio interés. Un agente económico es un ser humano cuyo comportamiento es inteligente y racional, y se considera que «el comportamiento de una persona es *racional* si actúa según *sus* mejores intereses, dada *su* información» (Aumann 2005: 351). Herbert Simon (1978: 2) argumenta que lo que caracteriza la racionalidad económica es su sentido maximizador individual, no sólo en las decisiones puntuales (aquí y ahora) sino también en lo concerniente a la formación de expectativas sobre ambientes de decisión, escenarios futuros y análisis estratégico de las decisiones. Robert Frank (1994: 227), por su parte, reconoce que en alusión al concepto de propio interés, «la mayoría de los economistas parte en sus investigaciones de alguna versión del criterio de la racionalidad basado en el egoísmo». Se entiende que el yo que calcula y decide en función de su propio interés es un sistema unificado de conciencia individual, integrado por un cuerpo físico y una mente rectora que toma decisiones sobre la base de una función de utilidad a maximizar. Para considerar a un ser humano

¹²⁶ Esta premisa es característica de la ciencia económica, y subyace al enfoque económico del comportamiento humano. Véanse Williamson (1993) y Becker (1997[1976b]).

como agente psicológicamente egoísta todas sus acciones han de estar motivadas por la consecución de algún tipo de beneficio subjetivo e intransferible (Sober y Wilson 2000: 198)¹²⁷.

Se manejan pues tres versiones teóricas de la racionalidad que pueden estar interrelacionadas teórica y metodológicamente: (a) la racionalidad como atributo exclusivamente humano, consistente en la elección consciente de comportamientos a través de operaciones de cálculo e inferencia; (b) racionalidad como utilitarismo, o requisito de una motivación basada en el logro de algún tipo de beneficio subjetivo; y (c) racionalidad como egoísmo psicológico.

Pasemos a contrastar estas hipótesis a la luz de la neurociencia contemporánea.

2.4.1.- De atributo exclusivo de la naturaleza humana a modelo de racionalidad instintiva

¿Es la racionalidad un atributo conductual sólo aplicable a los humanos, o también al resto de organismos? ¿Es la racionalidad un producto exclusivo del pensamiento humano, o un instinto o conglomerado de instintos implantados en todos los seres vivos? ¿El comportamiento inteligente emana de un cerebro que aprende y produce pensamientos racionales, o es una facultad independiente del pensamiento lógico-deliberativo?

Lionel Robbins (1935[1932]: 12 y ss.) se decantó por presentar una ciencia (económica) asignativa de recursos escasos, en la que el agente decisorio es el individuo, y además un *individuo humano* que, entre otros atributos psicológicos, tiene

¹²⁷ La representación del *homo oeconomicus* como un agente maximizador consecuencialista, reducido posteriormente a la expresión matemática de una curva de indiferencia (es decir, un conjunto coherente y ordenado de preferencias individuales), tiene su origen en la economía política decimonónica, y especialmente en los aportes teóricos y metodológicos de la economía marginalista. Con Jevons, y especialmente con la *Psicología matemática* (1881) de Edgeworth, triunfó la matematización deductiva del comportamiento humano, desligando la economía teórica posterior tanto de la psicología teórica y experimental, como de la neuropsicología que con el tiempo desembocará en las neurociencias de hoy día. No obstante, la psicología y la filosofía de la mente dominantes en los círculos académicos británicos durante la época en que se fraguó la Revolución marginalista sirvieron de soporte teórico, dado que el marginalismo británico, y antes el círculo intelectual empirista de los Mill, no fueron por libre de las corrientes de la psicología utilitarista. Véanse Davis (1995) y Giocoli (2003).

el de la conciencia de sí mismo («*sentient creature*»)¹²⁸. La economía asume pues la deriva antropocéntrica característica de las ciencias del comportamiento inspiradas en el racionalismo filosófico y su teoría de la elección: la elección racional¹²⁹.

Otras ciencias del comportamiento, como la psicología cognitiva, la etología y la psicología evolutiva, no han trazado tal línea demarcatoria entre la conducta humana y la del resto de animales de la naturaleza dado que, como subraya De Waal, «es cierto que los animales no son humanos, pero es igualmente cierto que los humanos son animales» (De Waal 2006: 65). Desde estos enfoques teóricos se parte de que existe una confusión extendida que consiste en considerar la inteligencia como un rasgo diferenciado y más elevado que los instintos. Los instintos, sin embargo, son racionales e inteligentes *per se*, en la medida que se trata de programas cognitivos adaptados y funcionales, aptos para resolver problemas de supervivencia¹³⁰. Lógicamente, la inteligencia es una característica propia de los programas cognitivos que permite organizar la información y concebir el mundo, y el propio organismo, de manera que éste consiga éxito reproductivo. Esta propiedad, que se encuentra en los mismos programas cognitivos y no en la toma consciente de decisiones, consiste igualmente en un programa cognitivo más (Loch, *et. al.* 2006: 223). Los instintos y programaciones epigenéticas de los organismos se muestran racionales e inteligentes, tanto en el análisis empírico como teórico:

Muchas especies calculan el tiempo que dedican a buscar alimento en distintos lugares, de modo que puedan optimizar su tasa de aporte de calorías por energía gastada en la búsqueda.

¹²⁸ A diferencia de Jevons y Edgeworth, Robbins tenía claro que la ciencia económica debía ser completamente independiente de la psicología. Por ello sustrajo a la economía de la incipiente neuropsicología (llamada biopsicología en su época), y estableció unos supuestos del comportamiento humano desde la propia economía separada de la biología:

El psicólogo debe explicar por qué alguien prefiere A a B, mientras que el economista sólo ha de asumir que estas diferencias existen... Él (el psicólogo) quiere saber por qué existen y a qué ley del equilibrio psíquico o genético obedecen. Nosotros (los economistas), por el contrario, necesitamos saber simplemente que existen para descubrir, *en nuestro propio campo y según las leyes de nuestra propia ciencia*, cuáles son las implicaciones de su existencia. (Robbins 1934: 99; paréntesis y cursivas añadidos).

La teoría de juegos, por su parte, ya nació con esta premisa de racionalidad despojada de psicología en los trabajos de Neumann y Morgenstern.

¹²⁹ Vernon Smith cita en esta visión de la mente humana a los filósofos, tanto racionalistas como empiristas, Descartes, Bacon y Hobbes; y a Bentham y John Stuart Mill, ya en el s. XIX, entre los filósofos ilustrados relacionados con la economía (Smith 2002: 504).

¹³⁰ Jeff Hawkins (2005) presenta la inteligencia como capacidad mental consistente en disponer de un modelo del mundo y hacer predicciones basadas en él. Esta versión de la inteligencia y de la racionalidad no debe excluir a los instintos y comportamientos inconscientes, así como a procesos decisorios estereotipados semiconscientes aprendidos en la socialización (Gigerenzer 2008).

Algunas aves aprenden la trayectoria del sol sobre el horizonte durante el día y a lo largo del año, información necesaria para navegar guiándose por el sol. La lechuza común utiliza discrepancias de un orden inferior a las milésimas de segundo entre los tiempos de llegada de un ruido a sus dos oídos para precipitarse sobre un ratón que se mueve en la hojarasca en plena oscuridad. Las especies que esconden alimentos colocan los frutos secos y las semillas en escondrijos que son impredecibles para así desbaratar los planes de los saqueadores, pero transcurridos varios meses tienen que recordar la posición de cada uno de los lugares y el cascanueces de Clark es capaz de recordar diez mil escondrijos. Los casos de manual que se aducen para ilustrar el aprendizaje por asociación, incluso los pavlovianos y los de condicionamiento operante, no resultan ser una retención de estímulos y respuestas coincidentes en el cerebro, sino algoritmos complejos para análisis de series temporales multivariantes y no estacionarias (que predicen que los sucesos ocurrirán, basándose en el historial de sucesos acaecidos). (Pinker 2004: 241).

Parte de la discrepancia entre ambos enfoques tiene un origen filosófico y antropocentrista, que radica en si se sitúa al ser humano en una posición diferenciada respecto al resto de seres vivos en la naturaleza o, por el contrario, se le da un tratamiento científico idéntico al de los demás seres vivos (Kanazawa 2004c). El atributo de racionalidad, como es el caso, se asocia implícitamente a un proceso lógico-deliberativo y de cálculo consciente, protagonizado por un yo-agente que conoce, evalúa y decide. Estas operaciones mentales sólo parecen poder ser realizadas por la especie humana, animal inteligente y racional, dotado de identidad personal, y por lo tanto un *ego* cuyas intenciones son hipotéticamente conscientes y causas de la acción. Sin embargo, la experiencia consciente sólo es una proporción muy reducida de las operaciones que se llevan a cabo en los módulos especializados del cerebro¹³¹. La mayoría de las complejas operaciones que realizan los seres humanos en su vida cotidiana se llevan a cabo mediante algoritmos no concebidos mediante la lógica formal o el cálculo racional, exactamente igual que ocurre en el resto de animales:

¿Una hormiga puede realmente calcular, hacer cálculos aritméticos? Desde luego, lisa y llanamente, no; pero tampoco lo hacemos nosotros cuando ejercemos nuestra facultad de navegar a estima, nuestro “sentido de la dirección”. Los cálculos de la navegación por integración de trayectoria se hacen de forma inconsciente, y su resultado asoma en nuestra conciencia -y, en el caso de que la tuviera, en la hormiga- como una sensación abstracta de que el lugar adonde vamos se halla en la dirección que seguimos, allí a lo lejos. (Pinker 2004: 240).

Así pues, la racionalidad no puede identificarse exclusivamente con la experiencia consciente y, dado que la especie humana comparte con los demás animales

¹³¹ Gazzaniga (1999: 44) precisa que alrededor del 98% de la actividad que se desarrolla en el cerebro no aflora a la experiencia consciente.

(y especialmente con los mamíferos) ciertos caracteres de la arquitectura y del diseño del cerebro y sus procesos, entonces, si la evolución «ha hecho bien su trabajo», cualquier ser vivo bien adaptado al medio ha de ser un sistema eficiente que “toma decisiones” adecuadas para su supervivencia (Dennett 1971: 93). Consecuentemente, como especifica Shizgal (2012: 5), los humanos comparten con los demás animales no humanos buena parte de los procesos emocionales y, en paralelo, las habilidades cognitivo-operativas, entre ellas la evaluación y planificación de las acciones, aunque los animales no humanos carezcan de ego. La conciencia reflexiva, por lo tanto, no es un requisito necesario del comportamiento “racional”, sino el producto de un cerebro más voluminoso y complejo, que sobre la base de un cerebro precedente ya dotado de innatismos racionales, dispone de recursos neurales adicionales para producir las innovaciones de un modelo de sí mismo y una teoría sobre la mente de los demás yoes (Shizgal 2012: 15). Existen numerosas pruebas de que esto es así, y la idea expresada en términos modernos y evolutivos tiene ya un siglo, como se explica a continuación.

El estudio experimental pionero de Edward Thorndike (1911) concluyó, desde una perspectiva de psicología comparada, en que los animales están provistos de inteligencia, y que su comportamiento es semejante al de los humanos ante elecciones e incentivos en situaciones decisorias comparables. Aunque un fenotipo semejante no tiene necesariamente que implicar una homología -puede resultar de una analogía, sin implicaciones filogenéticas-, sin embargo todos los seres vivos comparten el lejano ancestro común de la célula eucariota y, como ser vivo en un ambiente de escasez, tanto la célula eucariota como todos sus descendientes evolutivos han de estar adaptados para sobrevivir y propagar sus replicadores¹³². Este factor unificador es lo que nos permite comparar la racionalidad de los comportamientos humanos y no humanos. Por lo demás, admitimos que los cerebros de los organismos dotados de cerebro no son anatómica ni funcionalmente idénticos. Los vegetales, por su parte, carecen de cerebro. Sin embargo, puesto que todos los organismos comparten un mismo objetivo primario han de contar, si se encuentran bien adaptados, con sistemas productores de estrategias

¹³² En la entrada “Biological applications of economics”, de la edición de 1987 de *The New Palgrave*, Gordon Tullock (1987: 246) comenta la semejanza estructural que aprecia entre la biología y la economía, y afirma: «La hipótesis evolucionista en biología implica de manera contundente que los individuos de plantas y animales “actúan” como si intentasen maximizar la frecuencia de sus genes en el futuro».

de comportamiento equiparables, y con resultados racionales similares en situaciones similares¹³³.

Más recientes en el tiempo, los experimentos llevados a cabo con animales de laboratorio demuestran que los principales axiomas de la teoría económica del comportamiento del consumidor son extrapolables a la conducta sujeta a incentivos de otros animales (Kagel, *et. al.* 1995). Otras investigaciones, como las realizadas por Noë y Hammerstein (1994) y Noë, *et. al.* (1991) plantean la tesis de los “mercados biológicos” para explicar el comportamiento de diversas especies no humanas (aves, insectos y mamíferos) en su conducta de selección de “socios” (*partner choice*) a la hora de llevar a cabo procesos de reproducción, cooperación intra-específica y división del trabajo dentro de los grupos de una misma especie, así como mutualismo inter-específico. La principal conclusión de estos estudios consiste en que

... al igual que en el caso de los mercados humanos, la tasa de intercambio en los mercados biológicos está determinada por la ley de la oferta y la demanda. (Noë y Hammerstein 1994: 2).

Los animales estudiados operan como si de agentes racionales se tratara, y su comportamiento resulta bien explicado por modelos que siguen una estrategia decisoria que tiene en cuenta el coste de oportunidad:

... la ganancia futura esperada de la cooperación con un determinado socio es comparada con la que se obtendría de cooperar con otros socios potenciales, así como con la eficacia reproductiva esperada en ausencia de cooperación. (Noë y Hammerstein 1994: 3).

En juegos experimentales se han realizado ensayos con animales en los que éstos parecen elegir sus estrategias de cooperación o defección con criterios semejantes a los de un agente racional prototípico. Así, por ejemplo, Stevens y Stephens (2004) han dado continuidad a estudios previos sobre el comportamiento de arrendajos azules en estructuras de pagos diversas, y los resultados obtenidos son coherentes con los que se

¹³³ Esto conlleva aceptar que existe una explicación biológica del comportamiento económico, como analiza Robson (2001). Sin embargo, su trabajo sólo está referido a la especie humana, y no explica cómo operan los procesos neurológicos de causación del comportamiento económico. La neuroeconomía sí se está adentrando en este programa de investigación, aunque por el momento estudiando fundamentalmente el funcionamiento del cerebro humano. Véanse Camerer, *et. al.* (2005), y Glimcher (2003).

esperarían de un individuo humano que pretendiera maximizar los pagos. Esto, sin embargo, no sucede en elecciones de estrategia, dado que requieren facultades mentales avanzadas de las que no disponen los arrendajos, como el caso de proyección a futuro y descuento de resultados esperados¹³⁴.

Desde la década de 1950 existe evidencia experimental que demuestra que los animales no humanos tienen en cuenta los “precios relativos” en su comportamiento. Así, la recompensa relativa ofrecida a ciertos animales interviene empíricamente en su motivación y en su conducta, como si realizaran un análisis microeconómico de elección entre alternativas teniendo en cuenta los precios y la renta. Por ejemplo, David Zeaman (1949) entrenó a ratas para expresar sus preferencias por distintas cantidades de recompensa mediante la velocidad. Se pudo constatar una relación entre la velocidad a la que corrían y el nivel relativo de recompensa. Así, partiendo de una cantidad determinada, las ratas redujeron su velocidad cuando se les reducía la recompensa, y viceversa, partiendo de recompensas pequeñas, las ratas incrementaron su velocidad ante el incentivo de una ración mayor. En otros estudios con ratas, De Visser, *et. al.* (2011) concluyen que los dispositivos neurofisiológicos y sus respuestas conductuales asociadas son similares en la toma de decisiones de humanos y roedores, y ello tanto en individuos sanos como en casos de patologías neurológicas que impiden tomar decisiones según un planteamiento racional bayesiano.

Tremblay y Schultz (1999) han comprobado respuestas similares a las de las ratas en monos a los que se ofrecen alimentos que tienen una escala ascendente de preferencia en estos animales, como son los cereales, manzanas y uvas. Los monos distinguen perfectamente los tres tipos de recompensa y lo hacen expresando valoraciones relativas (disposiciones marginales a “pagar”) distintas ante la percepción de diferentes ofertas binarias de alimento. Manifiestan transitividad en sus preferencias, las ordenaciones son estables, son sensibles a incentivos, y la valoración revelada aparece claramente codificada no en una escala absoluta, sino en escala relativa. El nivel de excitación física que representa la “disposición a pagar”, y las valoraciones

¹³⁴ Ello se debe a que estas aves no disponen de las adaptaciones necesarias para la inferencia de intenciones de otros organismos, o para proyectarse a sí mismas hacia escenarios futuros alternativos al nivel que han desarrollado los humanos, es decir, no cuentan con una teoría de la mente.

relativas entre pares de alimentos están relacionadas, a su vez, con el nivel de actividad neuronal localizado en el córtex orbitofrontal del cerebro, un área que se sabe está especializada en la gestión de procesos de elección.

Otros experimentos llevados a cabo con plantas muestran resultados parecidos, en el sentido de que responden racionalmente a incentivos tales como los cambios inducidos en las condiciones ambientales (agua, nutrientes, luz), como si su comportamiento consistiera en la solución de un problema convencional de maximización sujeta a restricciones (Strong y Ray 1975)¹³⁵.

En todos estos casos se demuestra que los seres vivos se comportan según procesos racionales de gestión de la supervivencia, aunque no disponen de los atributos racionales tal y como los entendemos en la especie humana (lenguaje simbólico, pensamiento consciente, formulación de proposiciones lógicas, capacidad de cálculo e inferencia, etc.). Hoy por hoy se asume por la mayoría de la comunidad científica que sólo el *Homo sapiens* dispone de una conciencia subjetiva de sí mismo altamente desarrollada, y por lo tanto de la capacidad de tomar decisiones y hacer elecciones conscientes auto-atribuidas (Gazzaniga 2006: 21). No obstante, el que una acción sea llevada a cabo de manera refleja o instintiva, con la intencionalidad y la planificación de la conducta sumergida en el inconsciente, no implica que tal acción sea irracional. El diseño del sistema nervioso de la especie humana permite computar conscientemente una cantidad limitada de unidades de información por segundo que sólo representa el $4,5 \cdot 10^{-6}$ % del total de información que procesa¹³⁶. Obviamente, se trata de un diseño adaptado a la supervivencia, que economiza recursos como lo hace la mecánica de un automóvil utilitario en el tráfico urbano, o el metabolismo de un arbusto del desierto.

¹³⁵ Ello demuestra que no se requiere un cerebro para llevar a cabo conductas racionales, y que por lo tanto la racionalidad entendida como optimización no es un atributo del intelecto, sino un proceso de gestión de la vida. La vida en sí misma es ya optimizadora. Así, las adaptaciones llevadas a cabo en el reino vegetal no han requerido de un sistema nervioso que represente la diversidad y complejidad del mundo externo con el grado de riqueza informativa propio de las adaptaciones de los organismos dotados de movilidad. En el caso de los árboles y plantas ha sido suficiente la adaptación de tropismos de sus células especializadas (Llinás 2003: 18-20).

¹³⁶ El sistema nervioso humano puede computar 11 millones de bits por segundo de información aferente procesada a través de los sentidos externos. De toda esa cantidad de información, sólo 50 bits por segundo se procesan conscientemente (Nørretranders 1999: 125-126). Por lo tanto, la inmensa mayoría de la información y de las decisiones que “toma” el organismo quedan fuera de los procesos conscientes de cálculo y deliberación racional.

Esa economía de recursos es un indicio de racionalidad, y la capacidad de sobrevivir y reproducirse eficientemente es otro indicio, aunque los procesos decisorios sean inconscientes. Así pues, resulta pertinente tratar la racionalidad humana como un conglomerado flexible y complejo de instintos (Cosmides y Tooby 1994b). Darwin tenía claro que el comportamiento está organizado y motivado por los instintos, y que el comportamiento humano es una de las variedades instintivas de comportamiento existentes en la naturaleza:

El salvaje y el perro hallaron sin duda muchas veces agua a un nivel bajo, y la coincidencia de estas dos cosas llegó a asociarse en sus espíritus. Un hombre civilizado hubiera quizá formulado algún razonamiento general sobre esto; pero dado cuanto de los salvajes sabemos, nos hallamos autorizados a creer que probablemente no harían semejante razonamiento, y que el perro con seguridad no lo haría. A pesar de esto, tanto el salvaje como el perro, a pesar de los chascos que se lleven, buscarán siempre agua de la misma manera, y en uno y otro parecen estos actos hijos de la razón, tengan o no conciencia de que obran por un razonamiento. (Darwin 1972[1871]: I, iii. p. 76).

Las muestras de comportamiento económico racional en animales y plantas mediante procesos decisorios “rápidos y frugales” se basan, según Gigerenzer, *et. al.* (1999: 13) en una heurística no deliberativa, sino ecológicamente racional «en la medida que se adapta a la estructura de un entorno». La explicación que dan Kagel *et. al.* (1981b: 13-14) a los resultados “racionales” de los comportamientos de animales de laboratorio es una explicación darwiniana: la selección natural ha obrado de tal manera que tanto los humanos como los demás animales dispongan de repertorios de instintos que pueden ser caracterizados como soluciones a problemas de maximización sujeta a restricciones¹³⁷. Estos repertorios de instintos y otras conductas desarrolladas a través de aprendizaje, que motivan y determinan los comportamientos ante situaciones y estímulos de diversa índole, son patrones de respuesta hereditarios (epigenéticos) que necesariamente tienen, o tuvieron, valor de supervivencia. Sean los animales conscientes o no, y tengan o no la capacidad de realizar conscientemente cálculos e inferencias lógicas, todo ello resulta irrelevante, dado que su comportamiento es un producto evolutivo, y por lo tanto ha de ser coherente con una dinámica evolutiva adaptada para la supervivencia y reproducción (Plutchik 1980)¹³⁸.

¹³⁷ Téngase en cuenta que las plantas no han desarrollado instintos, sino tropismos.

¹³⁸ La clave está en los procesos de criba evolutiva de los comportamientos sub-óptimos, y no tanto en lo que sea determinado normativamente *a priori* como intencionalidad racional. Así, entidades carentes de motivación consciente, inteligencia o capacidad de aprender, acaban desarrollando comportamientos racionales (de maximización sujeta a restricciones) y generando equilibrios siempre

Pero entonces, ¿existe alguna diferencia en los procesos neurológicos subyacentes al comportamiento de un individuo humano respecto de los demás animales? ¿Están sometidos todos los seres vivos a patrones de acción completamente deterministas? La teoría de la elección racional asume implícitamente el principio de que existe un sujeto, un agente que fija objetivos, delibera para optimizar su consecución, elige entre objetivos y medios para su logro, y toma decisiones “libremente”. Para la neurociencia contemporánea lo que distingue a los humanos de otros animales es la arquitectura y diseño de su cerebro, lo cual influye en la posibilidad de que sean conscientes de *algunos* procesos decisorios. Todos ellos, conscientes o no, están causados no por la libre voluntad de un sujeto, sino por un complejo sistema de instintos interrelacionados. Asimismo, la neurociencia niega la neuroplasticidad estructural del cerebro, y por lo tanto las teorías clásicas del empirismo de la *tabula rasa*. Según la ciencia contemporánea la maleabilidad neural es un mito y una falacia. Sólo existen ajustes y refuerzos dinámicos, adaptativos al medio, en circuitos congénitos programados genéticamente para su desarrollo durante la vida del individuo, ya sea humano o no. Este desarrollo es completamente epigenético, como explica Michael Gazzaniga:

La infraestructura del cerebro humano está organizada al nacer. Sin duda su organización se debe a un sistema de crecimiento y retracción dinámico, que necesita procesos dependientes de la actividad a lo largo del desarrollo y de la vida posnatal. Los mecanismos genéticos de control desempeñan su papel en un medio físico-químico dinámico que conocen y con el que cuentan para ejecutar su tarea (...) Los mecanismos relacionados con las funciones mentales superiores (aprender a hablar o a razonar, por ejemplo) siguen siendo un misterio (...) Es obvio que el niño debe estar expuesto al lenguaje para adquirir los vocablos, pero la maduración cortical impulsa y restringe este sorprendente proceso (...) Los niños de dos meses empiezan a perder el reflejo de empuñar (la manera como el bebé se aferra a tu dedo cuando se lo pones en la mano). Además, experimentan una reducción de la sonrisa autogenerada, del llanto espontáneo, y mejoran su memoria de reconocimiento. ¿Resulta esto del aprendizaje? De ninguna manera. Cambios decisivos ocurren en la corteza. Muchos están bajo la égida del tallo cerebral, la porción más arcaica del cerebro, que controla las funciones motrices. Es probable que la expresión espontánea de reacciones como empuñar, sonreír o llorar se interrumpa cuando el control cortical invade el tallo cerebral a los dos meses de vida (...) Los circuitos cerebrales se están conectando según un plan y las modificaciones conductuales ocurren de modo automático en respuesta a los cambios de madurez del cerebro. Los circuitos cerebrales genéticamente programados están ejecutando su tarea. (Gazzaniga 1999: 86 y ss.)

que el sistema lleve a cabo una selección secuencial de reglas de comportamiento de acuerdo a restricciones de escasez (Gode y Sunder 1993). Por las mismas razones, organismos de esa misma naturaleza, caracterizados por una «profunda y debilitante ignorancia» llevan a cabo comportamientos aparentemente diseñados por una mente inteligente y racional. Esto se explica porque, como puntualiza Colman (2005: 217): «la evolución biológica con frecuencia imita el diseño inteligente».

¿Qué implicaciones metodológicas tiene para las ciencias sociales en general, y la economía en particular, el comprender e incorporar este modelo explicativo de la acción humana? Vernon Smith plantea que tal modelo alternativo al planteamiento clásico de la racionalidad es el que se corresponde con una racionalidad ecológica, y se basa en el hecho probado de que

... la acción humana está dominada por sistemas neurofisiológicos inconscientes y autónomos que permiten operar con eficacia sin hacer uso en todo momento del recurso cerebral más escaso: el funcionamiento de los circuitos relacionados con la atención y el razonamiento. (Smith 2002: 507).

2.4.2.- El protagonismo de las emociones y “demonios” en la elección racional

La neurociencia presenta la elección humana como un fenómeno psíquico ilusorio, igual que ocurre con el yo. Este fenómeno está integrado en los circuitos que regulan las emociones y otros demonios¹³⁹. También se integra en las habilidades y destrezas desarrolladas por condicionamiento, en su mayor parte por aprendizaje inconsciente, así como en la actividad motora del organismo¹⁴⁰. Gazzaniga explica que cuando la información relevante de un determinado estímulo o situación que requieren “decidir” aflora a la conciencia, en ese momento el cerebro ya ha empezado antes a operar activamente cursando instrucciones de conducta, o incluso ya ha realizado por completo un objetivo. Esto se lleva a cabo a través del despliegue, coordinación y aplicación de reglas automáticas dentro de procesos inconscientes, y ello ocurre a una velocidad vertiginosa para la capacidad humana de procesamiento consciente de información:

Cuando creemos que *sabemos* algo (es parte de nuestra experiencia consciente), el cerebro ya cumplió su tarea. Sin embargo, la noticia “nos” parece nueva. Casi ajenos a nuestra percepción consciente, los sistemas instalados en el cerebro trabajan por sí solos, automáticamente, y concluyen su trabajo medio segundo antes de que la información procesada alcance nuestra conciencia. En realidad, no es sorprendente que la mayor parte de

¹³⁹ La teoría computacional de la mente denomina *demón* a cada automatismo neurológico simple e independiente. Se trata de procesos instintivos que se activan ante la entrada de información en el sistema, y que producen a su vez información que activa otro/os demón/es (véase Pinker 2004: 102 y ss.) Incluso las operaciones más elevadas de la mente (como imaginar, inferir, empatizar, aprender o propagar conocimientos, hablar o calcular) están reguladas por demonios (Koch 2004).

¹⁴⁰ Un ejemplo de aprendizaje inconsciente es el del lenguaje y la argumentación lingüística: se trata de un instinto que se expresa y desarrolla ante la exposición a personas que hablan (Pinker 1991 y 1995). Los humanos además están genéticamente capacitados para un amplio campo de aprendizajes en la ontogénesis. Esto acaba marcando, junto con el etograma, el núcleo de la identidad personal y social del sujeto (Turner 2000: 152-153).

la actividad cerebral ocurra fuera de la conciencia; esta gran zona de actividad -donde se elaboran planes para hablar, escribir, jugar al tenis o levantar un plato de la mesa- funciona sin que siquiera sospechemos cómo lo hace. No planificamos ni articulamos estos actos: sólo observamos su rendimiento. Esta característica de la organización cerebro-mente vale en las percepciones más sencillas y en funciones más elevadas como la conducta espacial, las matemáticas e incluso el lenguaje. El cerebro disimula esta singularidad funcional creando la ilusión de que los sucesos están sucediendo en tiempo real y no *antes* del concurso de nuestra capacidad decisoria consciente. Muchos procesos que nos guían son actividades mentales, pero se asemejan a reflejos básicos en tanto que son adaptaciones preinstaladas, fabricadas por el cerebro cuando se enfrenta a un desafío (...) Aunque nuestro sentido de propósito y la centralidad de la voluntad aparezcan en primer plano, subyace en nosotros una maquinaria altamente especializada (...) Las técnicas de imágenes cerebrales nos permiten ver dónde y cómo actúa el cerebro antes de que surja una conducta: el cerebro decide antes. El yo consciente declara tomar una decisión que ya se ha procesado. (Gazzaniga 1999: 93-94, y 215).

¿Cómo funciona el cerebro cuando produce una decisión? La toma de decisiones es un complejo sistema de interrelaciones modulares que crea una sensación deliberativa guiada, en la que se percibe un dominio de la situación y una salida voluntaria de información dirigida al movimiento del cuerpo que llamamos comportamiento. La sensación de control autónomo es un espejismo motivado por tres factores que se analizan a continuación:

- a) La brecha temporal que existe entre el inicio de la elaboración de una respuesta motora y la conciencia de que se ha activado un proceso ejecutivo;
- b) la activación autónoma del sistema límbico, sin el cual no se produce decisión alguna; y
- c) el proceso de racionalización del proceso ejecutivo acometido.

2.4.2.1.- Tiempo y conciencia

El tálamo sensorial tarda alrededor de 500 milisegundos en captar estímulos somatosensoriales aferentes de alta frecuencia, como por ejemplo la picadura de un mosquito, para la producción de una experiencia sensorial consciente. El tiempo se reduce a 150 milisegundos para que la información comience a ser procesada inconscientemente¹⁴¹. Esa brecha temporal interviene en la explicación de la creencia

¹⁴¹ Durante ese tiempo el sistema de control y ejecución del cerebro ya ha empezado a trabajar con “hipótesis” y “reglas heurísticas” de origen epigenético, así como con creencias y reglas aprendidas y “grabadas” en el hipocampo y en otros circuitos neurales no congénitos. Véase Tononi y Edelman (1998: 1.848).

ilusoria en que el yo evalúa racionalmente la situación, y decide dar la orden a la mano de apartar al mosquito con un rápido movimiento¹⁴². Los experimentos llevados a cabo en 1983 y 1999 por Benjamin Libet y su equipo (Libet, *et. al.* 1983, y Libet 1999) midieron la cronología existente entre las variables *Voluntad subjetiva de realizar un movimiento* (V), *Momento en el que el sujeto es consciente de que ha comenzado el movimiento* (M), *Potencial de alerta activado en la corteza cerebral ocupada de las funciones motoras* (PA), y *Activación nerviosa de los músculos que producen el movimiento* (AM). El resultado de estos experimentos pone en tela de juicio la supuesta causación lógica de los movimientos en la que descansa la noción intuitiva de “decisión”: $V \rightarrow PA \rightarrow AM \rightarrow M$. Descubrieron que, como es de esperar, V antecede en el tiempo a M; pero hallaron también, y he aquí lo sorprendente, que PA se anticipaba en un intervalo de tiempo muy corto, pero positivo (300-500 milisegundos) a V. Esto permitió defender la tesis de que el cerebro prepara los movimientos que van a tener lugar en el cuerpo antes de que el individuo sea consciente de la voluntad de efectuarlos:

Este resultado induce a pensar que, en las personas, el sentimiento de intención puede ser un efecto de la actividad de preparación motora del cerebro, no una de sus causas. (Obhi y Haggard 2005: 16).

En 2003 algunos de los investigadores que colaboraron con Libet procedieron a repetir el experimento sobre la base de los anteriormente realizados (véase Wohlschlager, *et. al.*, 2003). En este caso se encontraron pruebas de que los procesos conscientes puedan ejercer algún efecto sobre la acción. Concretamente, y debido al desfase de unos 100 milisegundos entre el inicio de la preparación del movimiento y la orden de ejecutarlo, puede quedar margen para que la operación planificada aflore a la consciencia y que puedan generarse modificaciones de los diferentes planes motores que el cerebro ya ha puesto en marcha. De este modo, puede admitirse que la sensación

¹⁴² Esta ilusión está causada por un cerebro evolucionado, que sobre un sistema neural ancestral que funcionaba por completo con respuestas inconscientes, desarrolló un itinerario alternativo que posibilita la conciencia de algunos procesos ejecutivos y psicomotores. Así, el diseño primigenio presente en los homínidos es únicamente el integrado por el sistema límbico (circuitos aferentes \rightarrow tálamo sensorial \rightarrow amígdala), mientras que en el caso humano a este sistema se superpuso otro adicional evolutivamente reciente (circuitos aferentes \rightarrow tálamo sensorial \rightarrow córtex sensorial \rightarrow amígdala). El sistema dominante es el primitivo, y resulta mucho más rápido y eficaz en situaciones peligrosas que requieren una respuesta inmediata. El sistema moderno es comparativamente más lento, y funciona mejor en situaciones novedosas, en las que el organismo dispone de tiempo suficiente para reaccionar sin poner en peligro la vida. Véase LeDoux (2000).

de “intención” estaría relacionada con la elección de *cuál* será el movimiento que definitivamente se ejecute, y no con el proceso que activa el macroproceso de producir un movimiento para conseguir un objetivo. El individuo podría entonces inhibir o reconducir ciertas acciones ya iniciadas en los procesos motores, o bien elegir la más adecuada según las creencias y normas aprendidas¹⁴³. En cualquier caso, la polémica sigue abierta y todavía no se ha llegado a un amplio acuerdo científico sobre la naturaleza de los procesos de control y elección voluntarios. Tal y como explican Obhi y Haggard (2005: 20), el cerebro humano opera con modelos predictivos de representación de la realidad (incluido el propio organismo), y esos modelos procesan gran cantidad de información sensorial y memorística. La sensación de control y decisión surge de la consciencia de que el cerebro está llevando a cabo dichos cálculos en un momento dado, preparando una serie de alternativas de movimiento, y que exista en paralelo algún tipo de proceso ejecutivo que evalúe alternativas y seleccione la más conveniente de acuerdo al modelo que representa al organismo en interacción con otro modelo que representa el mundo externo.

El diseño del cerebro humano hace que primero se activen los procesos ejecutivos y después (milisegundos después) se produce la conciencia de tales procesos. Pero el cerebro no es capaz de producir la experiencia consciente de ese orden temporal, lo cual contribuye a reforzar la creencia, por propia experiencia, de que dichos procesos ejecutivos son un fruto de la voluntad y el razonamiento, cuando en realidad son de tipo acomodaticio: la volición y el razonamiento son una consecuencia y no la causa de la acción (Wegner 2003).

2.4.2.2.- Sin innatismos no hay elección racional

Lo que tradicionalmente se ha denominado “instintos”, haciendo alusión a los procesos orgánicos, pulsiones y acciones inconscientes, y por lo tanto aspectos del comportamiento no deliberado, hoy día son considerados por la ciencia como manifestaciones de un tipo de inteligencia pre-racional sin la cual resulta imposible tomar decisiones racionales¹⁴⁴. Estos procesos innatos, o innatismos, entre los que se

¹⁴³ Véase Edelman (1992: 143). Esto contribuiría a expandir las posibilidades de neuroplasticidad.

¹⁴⁴ Los trabajos de Tversky y Kahneman (1982; y 1987) han sido pioneros en la formalización y experimentación de modelos de comportamiento basados en reglas heurísticas y sesgos cognitivos de carácter epigenético. Véase también el enfoque de la inteligencia instintiva expuesto por Gigerenzer y

encuentran las emociones, son la base de toda elección racional y no sólo de la organización de las preferencias, dado que dominan todo el proceso neuropsicológico de la elección. Así, toda decisión comienza y finaliza en un cúmulo de emociones y sentimientos, y se gestiona en una cascada de procesos automáticos interrelacionados, conscientes e inconscientes (Damasio 2004; Evans 2001). La importancia de las emociones se sustancia en el hecho de que un robot ultrainteligente, con capacidades cognitivas extraordinarias, y con una base de datos enciclopédica para relacionarse con el mundo, carecería de motivaciones para actuar, y por lo tanto se vería atrapado en un bucle cognitivo similar al del famoso asno de Buridan (Dennett 1984). Seleccionar objetivos y alternativas de elección ante una determinada tesitura requiere una previa valoración relativa de todo ello, y esa valoración se lleva a cabo mediante las emociones. De este modo,

.... ser racional significa ser emocional, y cualquier acción que escojamos habiendo separado cognición y emoción, fallará al hacerlo el proceso neurológico de la cognición. (Turner 2000: 59).

La diferencia existente entre animales humanos y no humanos no es una cuestión de racionalidad *versus* instinto, sino de que el género humano dispone de una cantidad muy superior y más variada de instintos (todos ellos racionales) que permiten concebir una cantidad extraordinariamente superior de valoraciones emocionales y alternativas de elección (Pinker 2004: 244-245).

¿Cómo participa el sistema emocional cuando el organismo toma una decisión? Todo organismo dispone de un sistema de evaluación de su homeostasis y de la interrelación de ésta con el entorno. El procesamiento de datos no es suficiente para la evaluación, puesto que los datos han de ir acompañados de una valoración asociada a la homeostasis en tiempo real. En el sistema nervioso humano la valoración se lleva a cabo de manera autónoma en el sistema límbico, que es el dispositivo encargado de gestionar las emociones. Sin ellas, los procesos ejecutivos no pueden activarse porque no “comprenden” el significado de la información para el organismo (Damasio, *et. al.* 1991). Estas operaciones no son realizadas por las estructuras neurales encargadas del pensamiento lógico, el cálculo y el lenguaje (los lóbulos frontales), sino por estructuras

sus colaboradores (1999; y 2000) y por Kahneman, *et. al.* (2002) en otra línea de investigación.

más primitivas pertenecientes al sistema límbico. Así, aunque éstas funciones mentales conscientes son las más recientes en la filogénesis, sin embargo: «Nada dramático le ocurre a la conciencia si son dañados los lóbulos frontales» (Ramachandran y Hirstein 1997: 450). No ocurre lo mismo en el caso de daños en el sistema límbico o en las conexiones entre éste y los subsistemas encargados de gestionar los procesos de planificación y ejecución de las acciones. Problemas de este tipo incapacitan al individuo para tomar decisiones. Damasio cuenta la anécdota de uno de sus pacientes, aquejado de una lesión en el lóbulo temporal que le había producido la desconexión entre los módulos afectivo y cognitivo de su cerebro. Un día le preguntó que cuándo prefería quedar con él para la siguiente sesión terapéutica, ofreciéndole como alternativas dos fechas posibles separadas tan sólo por unos días:

Durante casi media hora el paciente enumeró razones a favor y en contra de cada fecha: compromisos previos, proximidad de otros compromisos, condiciones meteorológicas, casi cualquier cosa que se pudiera razonablemente ligar a una simple fecha... Nos hizo aguantar un tedioso análisis de costes y beneficios, así como una interminable descripción y una inútil comparación de opciones y posibles consecuencias. (Damasio 2004: 193).

Tras un largo y lógicamente coherente discurso racional, al final el paciente aceptó sin más la fecha que el neurólogo le propuso.

La conciencia tiene su sede principal en los procesos neurológicos que ocurren en las amígdalas y otras áreas situadas en los lóbulos temporales, un subsistema del cerebro que hace de interfaz entre la percepción y la acción consciente. Así, las experiencias conscientes son el puente que une las etapas iniciales del proceso perceptivo y las etapas finales previas a la acción, en las que los procesos ejecutivos del cerebro despliegan objetivos y alternativas de comportamiento, la planificación y la ejecución motora más adecuada a las circunstancias¹⁴⁵.

¹⁴⁵ La etapa final del proceso que dictamina el comportamiento a seguir, en la que el cerebro lleva a cabo valoraciones y selecciona la mejor estrategia de acción, se desarrolla dentro del sistema límbico (concretamente las amígdalas y el giro cingulado anterior), que es donde residen los procesos ejecutivos (y no en las zonas frontales y prefrontales). En el proceso ejecutivo se conectan los *qualia* perceptivos con emociones y propósitos concretos, lo que permite que se produzcan decisiones (Ramachandran y Blakeslee 1999: 314-315). Como ya hemos indicado, la falta de datos sensoriales para confeccionar representaciones fiables y ricas en información es un problema de puntos ciegos perceptivos. Es por esto que el cerebro rellena los *qualia* insuficientes o defectuosos de forma automática, y lo hace en esta fase de proceso de control, en la que se completan los vacíos de la mejor forma posible a fin de que las estructuras límbicas puedan trabajar con una información completa y fiable, y así seleccionar las respuestas motoras y de planificación más adecuadas para la supervivencia.

Lo que ocurre con los pacientes incapaces de decidir, o que objetivamente toman decisiones disparatadas o incoherentes, es que su cerebro representa la información de forma anómala, y/o no es capaz de asignar significados a los *qualia* (están incapacitados para una genuina semántica emocional)¹⁴⁶. Tales significados no son de orden lógico, sino esencialmente emocionales, y la elección entre alternativas necesita no sólo de la emoción *in situ*, sino que además requiere de la memoria emocional autobiográfica, que permite conectar las experiencias pasadas y su significado emocional con la acción presente. Muchas de estas patologías son incompatibles con la supervivencia autónoma del individuo enfermo, puesto que esa semántica emocional a la que aludimos proporciona la información y las señales oportunas, la motivación necesaria, para que el cerebro gestione las respuestas acertadas para la supervivencia y la replicación. Esas señales, que hoy día se denominan *marcadores somáticos*, se emiten en un lenguaje emocional (bienestar/malestar; positivo/negativo; atracción/repulsión) y esas intuiciones, aparentemente carentes de justificación ni discurso lógico, resultan poseer la “lógica paradójica” de la evolución, con frecuencia inaccesible al entendimiento racional del individuo, pero sin embargo adaptativa y eficaz para sobrevivir y replicarse con éxito:

En este contexto, el término racional no denota un razonamiento lógico explícito, sino más bien una asociación con acciones o resultados que son beneficiosos para el organismo que muestra emociones. Las señales emocionales recordadas (los marcadores somáticos) no son racionales en y por sí mismas, pero promueven consecuencias que podrían haberse obtenido racionalmente. (Damasio 2005: 146).

Damasio ha demostrado, mediante experimentos de emoción inducida y técnicas de medición de la respuesta galvánica de la piel, que la cognición de estados mentales que se asocian a sentimientos (como la alegría y la tristeza) viene precedida sistemáticamente por estados del cuerpo y mapas neurales diferenciados¹⁴⁷:

Habíamos supervisado continuamente las respuestas fisiológicas de los sujetos y pudimos advertir que los cambios en conductancia dérmica *precedían siempre* a la señal de que se

¹⁴⁶ Téngase en cuenta que los *qualia* son representaciones mentales en bruto, carentes de significado emocional, como lo doloroso del dolor, la rojura del rojo, el calor de lo caliente, o lo agradable del bienestar (Ramachandran y Hirstein 1997: 430n₁).

¹⁴⁷ La respuesta dermogalvánica consiste en la transpiración de la piel causada por emociones que resultan activadas ante la percepción inmediata de un suceso. Se trata de que el sudor cambia la conductividad eléctrica de la epidermis, y existen técnicas de laboratorio que permiten medir este fenómeno.

estaba sintiendo un sentimiento. En otras palabras, los monitores eléctricos registraron la actividad sísmica de la emoción, de manera inequívoca, *antes* de que los sujetos movieran su mano para indicar que la experiencia había empezado (...) el experimento ofreció pruebas adicionales de que los estados emocionales llegan primero y los sentimientos después. (Damasio 2005: 100).

Este experimento demuestra que antes de que el sujeto sometido a una simulación de riesgo sea capaz de entenderlo lógicamente y concentrarse en las posibles decisiones, la respuesta dermogalvánica ya se ha producido con antelación, como si el cerebro “intuyera” el peligro y diera pistas emocionales (todo ello automática e inconscientemente) que guían la acción consciente a través de una supuesta “elección racional”¹⁴⁸.

2.4.2.3.- Racionalidad y racionalización

Siendo el comportamiento causado por procesos funcionales autónomos del cerebro y no por decretos decisorios de un supuesto agente interno, ¿por qué los individuos neurológicamente sanos tienen la impresión de que dirigen voluntariamente sus actos y toman decisiones acerca de su comportamiento? George Loewenstein (1996: 276) afirma que

... la mayoría de las personas experimenta sus propias acciones como fruto del resultado de sus decisiones, o al menos como previamente deliberadas. Sin embargo, puede cuestionarse si esas introspecciones representan verídicamente los procesos de decisión subyacentes, o se trata más bien de racionalizaciones *ex post* del comportamiento.

En este sentido, Carlsmith (2008) ha demostrado experimentalmente que un porcentaje mayoritario de personas muestran una actitud de aprobación del contenido de normas legales abstractas que, cuando se llevan a la práctica y se aplican, rechazan por considerarlas injustas. Según el autor, esto se debe a su incapacidad para conocer sus propias motivaciones. Existe un conflicto entre la valoración emocional y cognitiva de las normas. La valoración emocional suele anteceder a la cognitiva, y esto provoca ulteriores esfuerzos racionalizadores de la discrepancia.

¹⁴⁸ Coricelli, *et. al.* (2010) han comprobado experimentalmente que las emociones desagradables de la culpa y del miedo se manifiestan a nivel inconsciente y somático en la respuesta dermogalvánica de la piel antes de que se tomen conscientemente decisiones arriesgadas, como por ejemplo, la de evadir impuestos. La respuesta somática inconsciente se incrementa en intensidad conforme aumenta el valor del riesgo asumido, y especialmente si se incluye una sanción consistente en la publicidad de una fotografía del evasor.

El cerebro humano está estructurado en demonios y módulos especializados que llevan a cabo funciones de manera automática, y sólo una pequeña parte de los procesos gestionados (alrededor del 2%) produce experiencia consciente cognoscible y expresable mediante el lenguaje. En el *Tractatus*, Wittgenstein escribió que: «Lo que no podemos pensar no podemos pensarlo. Tampoco, pues, podemos decir lo que no podemos pensar» (Wittgenstein 1999[1921]: §5.61). Efectivamente, el cerebro está compuesto de dos hemisferios que integran subsistemas modulares con multitud de funciones diferenciadas (véase Figura 2.4.) El hemisferio izquierdo, que es el dominante en las personas diestras, está especializado en la producción del lenguaje, el razonamiento, el cálculo y la resolución lógica de problemas. El derecho está especializado en tareas de organización de la percepción en un modelo tridimensional de percepción espacial, así como en otros aspectos relativos al reconocimiento de rostros, la imaginación y la creatividad. Las acciones generadas en procesos neurológicos que no están conectados con el subsistema ubicado en el hemisferio izquierdo tienen su propia intencionalidad inconsciente, pero el hemisferio dominante, que dispone de un subsistema de racionalización del comportamiento, inmediatamente emite teorías y justificaciones de las acciones realizadas; por eso Gazzaniga (1993) denomina a ese módulo el “intérprete”. Obviamente, las acciones inconscientes realizadas que el intérprete no pueda justificar y argumentar no pueden ser pensadas, y, como apuntó Wittgenstein, por ello tampoco pueden ser expresadas a través del lenguaje.

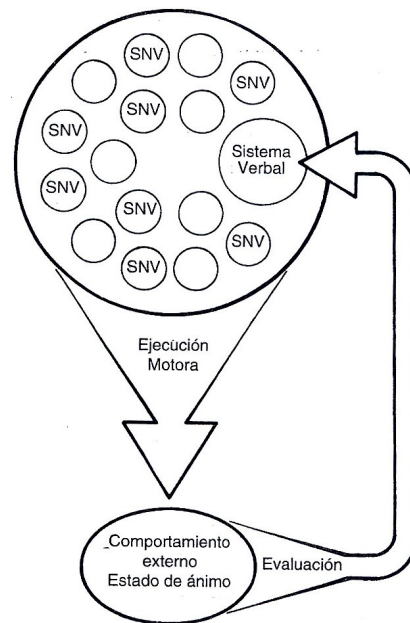


Figura 2.4.- El cerebro está organizado en múltiples sistemas modulares (cientos, o incluso miles) de procesamiento, que se manifiestan en acciones del cuerpo y sentimientos, y no por medio de la comunicación verbal. Estos sistemas producen actividades sin intervención del lenguaje y sin conciencia de que ocurren. Una vez producidos los comportamientos o los sentimientos, éstos se convierten en la información objetiva con la que el subsistema lógico-lingüístico, ubicado en el hemisferio izquierdo, produce interpretaciones y teorías acerca del significado de los comportamientos y sentimientos. (SNV, son las siglas de *Sistema No Verbal*).

Fuente: Gazzaniga (1993: 117).

En casos de pacientes con enfermedades neurológicas que han sido terapéuticamente hemiseccionados, la información presentada a través de los sentidos de la vista o del tacto por los canales aferentes de cada hemisferio produce comportamientos a los que el módulo intérprete provee de narrativas racionalizadoras diferentes. Así, cuando a uno de estos pacientes se le proporciona una información a través de los canales aferentes del hemisferio derecho, y esta información activa procesos emocionales y ejecutivos que desembocan en acciones y estados de ánimo, el módulo intérprete, que no ha participado de la información originaria,

... urde la hebra que vincula los sucesos. Dicho de otro modo, consigue que las acciones y estados de ánimo parezcan dirigidos, articulados e intencionales. (Gazzaniga 1999: 174).

Valga un ejemplo: se proyecta únicamente al hemisferio derecho la orden “camina”. La respuesta del paciente es levantarse de la silla y disponerse a caminar. Si se le pregunta

que adónde va, o por qué se ha levantado, las contestaciones suelen ser pintorescas, como por ejemplo: «voy a coger una Coca Cola a mi casa», «me duele la espalda de estar tanto tiempo sentado», etc. En todos estos casos, Gazzaniga justifica que «el sistema cognitivo del cerebro izquierdo necesita una teoría, e instantáneamente crea una que, dada la información que tiene sobre esa tarea concreta, tenga sentido» (Gazzaniga 1993: 108).

Cabe suponer que este “instinto de racionalización” haya sido implantado por la evolución para evitar lo que en psicología se denomina disonancia cognitiva, es decir, la discrepancia entre lo que se cree, lo que se dice y lo que se hace. Un cerebro tan complejo como el humano, que dispone de modelos de sí mismo y del mundo externo, y que está capacitado para una conciencia autorreflexiva que maneja creencias y teorías, no admite incoherencias dentro del propio sistema, pues de ser así el sistema carecería de eficacia, especialmente en las relaciones con el mundo y la interacción social. Por eso la coherencia se impone *a posteriori* para reforzar la estabilidad teórica del sistema y para reforzar el sentido de control de la situación y en definitiva el yo ejecutivo. Este proceso contribuye a proporcionar la certeza de existencia de un sujeto¹⁴⁹.

Lo que se llama “decisión” tiene que ver con una rápida interpretación retrospectiva del acontecer psicomotor, de modo que la “voluntad de hacer algo” vendría a ser una inferencia del momento en que el cerebro emite una orden de actuar, y no una intuición causal o una causa directa (Banks e Isham 2009). El razonamiento, cuando tiene lugar, encuentra explicaciones a los actos y, aunque esas explicaciones son *a posteriori*, existe la ilusión de que son *a priori*. No es de extrañar que Descartes creara el mito del fantasma en la máquina impresionado por esta ilusión, sin embargo: «nuestra mente consciente es más un observador que un agente decisorio» (Trivers 2000: 120).

¹⁴⁹ Nótese que la experiencia subjetiva del *yo*, y no digamos la visión histórico-narrativa de la identidad personal, se asemejan a una operación inconsciente de relleno de puntos ciegos supuestamente autobiográficos. La neuropsicología considera que el cerebro confabula para ofrecer la experiencia de una auto-imagen coherente y estable, para lo cual distorsiona y retoca los recuerdos, y rellena vacíos de memoria (Eakin 2001).

2.4.3.- Refutación de la racionalidad del egoísmo optimizador

El juego del dilema del prisionero ha dado lugar a gran cantidad de literatura académica sobre la cooperación. Prescribiendo la acción humana dentro de este juego, y en jugada única (*single spot game*), la decisión de actuar cooperativamente carece de sentido racional (Nash 1950 y 1951). Esto es así ya sea en la decisión inicial, o en la respuesta a un comportamiento cooperativo previo por parte de uno de los jugadores. La estrategia no-cooperativa es la predicción de lo que Ostrom (2000: 140) denomina “modelo estándar de acción racional individual”, o modelo del “egoísmo racional”. Según estos supuestos, el equilibrio de Nash es un cruce de decisiones no-cooperativas en el dilema del prisionero de una única jugada. Sin embargo, no pocos han sido los alegatos en la literatura especializada favorables a la racionalidad de la cooperación en este contexto, muchos de ellos falaces o incluso bienintencionadamente ingenuos¹⁵⁰. El caso es que, aun admitiendo la debilidad de muchas de estas hipótesis, y a pesar del pretendido determinismo teórico de la elección racional, la experiencia personal acredita que en más ocasiones de lo que pudiera parecer los individuos humanos actúan de manera altruista, cooperan frecuentemente incluso en jugadas anónimas de una sola tirada en juegos cuya estructura coincide con la del dilema del prisionero, y en la práctica eso aporta ganancias reales superiores a las que se derivarían de la elección racional del modelo estándar (Colman 2003).

Muchas personas cooperan en intercambios de mercado con desconocidos a quienes creen que no volverán a ver, pudiendo engañarles y obtener ganancia de ello. Asimismo, muchas personas revelan un genuino interés por otros, realizando acciones que suponen importantes sacrificios que no tendrían por qué asumir desde la perspectiva de la elección racional. Por ejemplo,

En el mundo jurídico se sabe que normalmente las personas anteponen los intereses de sus familias, amigos, y otros grupos a los que pertenecen, a sus propios intereses individuales. (Colman, *et. al.* 2008b: 409).

Experimentalmente, y en trabajos de campo realizados tanto dentro como fuera del laboratorio, se aprecia una clara tendencia en las poblaciones de individuos a llevar a

¹⁵⁰ Véase Binmore (1994, cap. 3).

cabo comportamientos prosociales como puras externalidades positivas, es decir, transferencias de recursos al margen del mercado, como por ejemplo contribuir con donaciones no condicionadas a entidades benéficas en circunstancias completamente anónimas (Meier 2007; Frey y Meier 2004a).

La extensión e intensidad del comportamiento prosocial varía de forma significativa entre individuos. Así, mientras que algunos se comportan en dilemas sociales de manera canónica de acuerdo a la teoría del actor racional, otros revelan claramente que les preocupan y ocupan los intereses de otras personas, y lo que podríamos llamar las normas sociales del bien común (Ostrom 2000: 138). La psicología social ha clasificado a los humanos en función del carácter más o menos social de sus preferencias en grupos de comportamiento; por ejemplo: (a) individualista, competitivo y cooperativo (véase p.ej. McClintock 1978); y (b) cooperativo y competitivo (Kelley y Stahelski 1970). Hoy día se está empezando a conocer la posible base epigenética del comportamiento, tanto prosocial como antisocial (Hare 1984; Lykken 1995; Kendler, *et. al.* 2007; Morishima, *et. al.* 2012). Los hallazgos de la economía experimental corroboran esa segmentación de la población en individuos de disposición tendente a motivaciones egoístas o prosociales, siendo las proporciones respectivas más o menos similares (en torno al 50% cada grupo) (Andreoni y Vesterlund 2001; Charness y Rabin 2002; Weimann 1994).

Los estudios experimentales llevados a cabo mediante juegos interactivos han demostrado que en proporciones nada desdeñables (desde el 40%, pasando por el 50%-60%, y hasta más del 80%) los agentes empíricos cooperan en situaciones del tipo dilema del prisionero, observándose que la tasa de cooperación se mantiene elevada si los jugadores interaccionan repetidamente y el final de las repeticiones queda lejos o es indeterminado¹⁵¹. Incluso en juegos de jugada única se aprecia una elevada tasa de cooperación (42% - 50%); y lo que resulta aún más insólito: en iteraciones finitas, en las que los jugadores han podido aprender y afinar sus hipotéticas estrategias de

¹⁵¹ A este respecto pueden citarse los trabajos pioneros recogidos en el meta-análisis de Sally (1995), que recoge conclusiones de experimentos llevados a cabo entre 1958 y 1992, así como el experimento más largo publicado en la literatura especializada (Rapoport y Chammah 1965). Otros estudios experimentales que pueden citarse al respecto son los de Colman (1995 y 2003); Clark y Sefton (2001); Good (1991); Rapoport (1989: cap. 12), y Grzelak (1988).

cooperación condicional con la experiencia acumulada, la evidencia experimental muestra que, aunque menor, la tasa de cooperación resultante es contraintuitivamente bastante alta (superior al 30%). En las últimas jugadas, tras una cantidad a veces muy numerosa de iteraciones previas, un porcentaje significativo de jugadores (entre el 10% y el 20%) coopera contra pronóstico (Andreoni y Miller 1993, Cooper, *et. al.* 1996, y Selten y Stoecker 1986).

Las versiones del juego del dilema del prisionero planteadas desde una perspectiva evolucionista en modelos informáticos de inteligencia artificial revelan también altas tasas de cooperación (Axelrod 1997: caps. 1 y 2; Kraines y Kraines 1995; Nowak y Sigmund 1993a y 1993b). Más de la mitad de los jugadores cooperan en el dilema del prisionero con contrincantes humanos, y algo menos incluso lo hacen en interacciones con programas informáticos (Rilling, *et. al.*, 2004a). Cuando se pregunta a los jugadores que siguen conductas cooperativas *a priori* contrarias a la lógica de la elección racional por qué actúan así, ofrecen con frecuencia justificaciones de tipo moral, como reglas de comportamiento recogidas en las religiones o simplemente la voluntad de mantener la conciencia tranquila y eludir el sentimiento de culpa. Indagando aún más afloran motivaciones relativas al sentido de pertenencia grupal, el conocimiento mutuo y la vinculación, la empatía y simpatía, y la actitud de tolerancia y apertura básicas para convivir en paz (véase Poundstone 1995: 333 y ss.)

En experimentos llevados a cabo con otros dilemas sociales, como el juego del ultimátum, el juego del dictador y los juegos de bienes públicos también aparecen resultados contraintuitivos para el modelo estándar de racionalidad.

En el juego del ultimátum se aprecia en todos los estudios realizados, con personas de distintas características sociodemográficas y culturales, que se tiende a proponer un reparto más o menos equitativo (y a rechazar repartos excesivamente desequilibrados)¹⁵². Así, los resultados muestran que repartos en torno al 50% tienden a ser mayoritariamente aceptados, mientras que otras propuestas de reparto menos

¹⁵² Los trabajos pioneros son de Güth, *et. al.* (1982). Thaler (1992: cap. 3) realiza un amplio repaso de los experimentos con el juego del ultimátum, al igual que Oosterbeek, *et. al.* (2004) con un meta-análisis basado en los resultados de 50 trabajos académicos.

equilibradas (más allá del límite 80% - 20%) son rechazadas de forma sistemática (Henrich, *et. al.* 2005; Samuelson 2005). La tasa de aceptación de propuestas de reparto alejadas del equilibrio (50% - 50%) decrece en mayor medida conforme aumenta la desigualdad de la oferta realizada¹⁵³.

En las diversas variantes del juego del dictador los resultados concluyen en que el dictador no se queda con todo (que sería lo más fácil y beneficioso para él, dado que no puede recibir represalia por parte del receptor ni arriesgar su reputación ante terceros), sino que la distribución más frecuente es una donación situada entre el 20-30% del recurso valioso a repartir (véase Hoffman, *et. al.* 1996)¹⁵⁴. Además Hoffman, *et. al.* (1996), Brañas-Garza (2006 y 2007), y Brañas-Garza, *et. al.* (2009) encuentran que, partiendo del completo anonimato (en el que las donaciones son mayoritariamente positivas) la tasa de reparto se incrementa conforme aumenta el personalismo relacional con el receptor, la información acerca de la situación personal del mismo, así como el carácter y la reputación de la persona o entidad receptora (por ejemplo, si se trata de una entidad como la Cruz Roja u otras organizaciones dedicadas a ayudar a personas necesitadas).

En cuanto al análisis de los resultados con juegos de provisión y sostenimiento de bienes públicos, se llega a la conclusión de que también los jugadores cooperan en niveles sorprendentemente elevados, tanto más cuanto menor sea el tamaño del grupo, mayor la proporción de individuos que penalizan a los aprovechados (*strong reciprocators*), existencia de un resolutivo liderazgo pro-cooperativo, mayor eficacia de los sistemas de información, vigilancia y detección del *free-riding*, mayor el coste que suponen los castigos a los defectores, mayor la proporción de cooperadores condicionales, mayor la homogeneidad de rasgos biológicos y étnico-culturales, mayor

¹⁵³ Las tasas respectivas de rechazo de ofertas no equilibradas vienen a ser las siguientes: un reparto 7:3 es rechazado por el 5% de los agentes; 8:2 por el 47%; y 9:1 por el 61% (véase Sanfey, *et. al.* 2003).

¹⁵⁴ Téngase en cuenta que este experimento consiste en que se le entrega una cantidad a un individuo (el dictador) indicándole que tiene la libertad de distribuirla como quiera entre él y otra persona que no tiene capacidad de responder *a posteriori*, ni de intervenir estratégicamente *a priori*. La información y el grado de conocimiento sobre las identidades respectivas pueden variar según se lleve a cabo el marco cognitivo (*frame*) del experimento. En estos casos tal información puede afectar a la decisión del individuo proponente, y ese aspecto es un objeto habitual de análisis en los experimentos.

la estabilidad grupal, y más dominantes las preferencias individuales por la inclusión social en el grupo¹⁵⁵.

Obsérvese que tanto en el juego del ultimátum, como en los de provisión y sostenimiento de bienes públicos, la actitud de castigo (consistente en la venganza y represalia por parte del receptor directo, o por parte de terceros) desempeña un papel relevante en el resultado de la asignación. Los resultados experimentales indican que son muy frecuentes los comportamientos de castigo, incluso cuando el mismo castigo resulta oneroso para quien lo aplica, e incluso cuando quien lo ejecuta no es parte directamente agraviada (Fehr y Schmidt 1999; Fehr y Gächter 2000b). Este comportamiento también resulta contraintuitivo desde la lógica de la elección racional, y aunque se puede alegar que se trata de maximizar la valencia hedónica de ciertas experiencias psicológicas (el alivio de una tensión emocional causada por emociones tales como la ira o el resentimiento), sin embargo la realidad es que el sujeto gasta efectivamente recursos (paga un precio) para castigar personalmente a otro, cuando podría disfrutar de esa utilidad externalizando el coste de aplicar el castigo, y dejando así que sean otros individuos quienes costeen en precio de la represalia.

2.4.4.- Un utilitarismo matizado y una función de utilidad compleja: la racionalidad de la elección sin raciocinio

La acepción de la racionalidad entendida como utilitarismo no atiende las motivaciones intrínsecas, sino que parte del supuesto de que toda motivación está orientada a una ganancia neta de “algo”. De este modo, toda acción es instrumental, un medio para conseguir algo, y nunca un fin en sí mismo (salvo que ese “algo” se halle implícito en la ejecución de la acción). Esta viene a ser la síntesis más divulgada de los modelos microeconómicos del comportamiento. En microeconomía el agente racional maximiza utilidad en función de unas preferencias (deseos), y teniendo en cuenta el cálculo de probabilidades (creencias) (Robson 2002: 89). Los agentes económicos

¹⁵⁵ Una muestra de la literatura respectiva es: Ledyard (1995); Sally (1995); Ostrom (1998); Van Lange, *et. al.* (1992); Brewer y Kramer (1986); Isaac y Walker (1988); Fehr y Gächter (2000a y 2002); Fischbacher, *et. al.* (2001); De Cremer y Leonardelli (2003); Sausgruber (2009); Frey y Meier (2004b); Mittone y Ploner (2010); Rustagi, *et. al.* (2010).

actúan como si trataran de maximizar el valor esperado de algo, y ese “algo” es lo que la teoría de las preferencias reveladas denomina utilidad (Binmore 2007: 5)¹⁵⁶.

La función de utilidad empírica es bastante más compleja que la teorizada por los economistas decimonónicos. Existen numerosos trabajos de economía experimental y neuroeconomía que demuestran que la obtención de recursos materiales no es una buena (ni la única) *proxy* de la utilidad, esto es, de lo que valoran las personas empíricas. Una teoría sobre la función de utilidad no puede obviar, como ya hemos indicado, que la mayoría de los participantes en experimentos rechazan un reparto inferior al 20% en juegos del ultimátum, aunque la aceptación maximizaría la riqueza (Oosterbeek y van de Kuilen 2004; Samuelson 2005). Muchas personas practican costosos castigos altruistas a los oportunistas en juegos de bienes públicos (Fehr y Gächter 2000a; 2000b, y 2002); regalan dinero con la misma satisfacción que les produce recibirlo regalado (Moll, *et. al.*, 2006); experimentan satisfacción haciendo regalos y contribuyendo a mejorar la situación económica de otras personas (Camerer 2003; Fehr, *et. al.* 2005); y, en general, existen procesos automáticos de respuesta emocional que anteponen la justicia, la reputación, la generosidad y la solidaridad al mero enriquecimiento personal (Gintis, *et. al.*, 2005; Jones y Goldsmith 2005; Frank 1994: cap. 7; Frank 1988). Incluso se han llevado a cabo experimentos en los que se demuestra que hay personas (y otros animales no humanos) que realizan “castigos perversos” con un alto coste personal¹⁵⁷. La psicología humana muestra paradojas que no por ello dejan de ser coherentes con una perspectiva racional del comportamiento que no es el utilitarismo de la riqueza material:

¹⁵⁶ Las dos acepciones de la racionalidad utilitaria, la de la microeconomía y de la teoría de juegos, se representan en lo que Paramio (2000: 67) denomina *racionalidad paramétrica* y *racionalidad estratégica*. La primera, característica de la microeconomía, supone al individuo actuando siempre en un mercado, provisto de recursos y de una función de utilidad con preferencias ordenadas y no contradictorias, e información relevante para maximizar su utilidad. La segunda variante, que corresponde al estilo decisorio de la teoría de juegos, no supone al individuo actuando en un mercado, sino ante al menos otro individuo, e interaccionando dentro de un marco de reglas del juego. Cada individuo no sólo se ve abocado a calibrar las repercusiones de su conducta en su función de utilidad (que igualmente pretende maximizar), sino que debe considerar las probables decisiones que tomen los demás intervinientes en la interacción, y la repercusión de éstas en sus resultados individuales.

¹⁵⁷ Un castigo perverso consiste en producir daño sin obtener un beneficio por ello, algo así como comportamientos mezquinos, envidiosos o malévolos (*spiteful behaviors*). Véase al respecto Hamilton (1970). Asimismo, para experimentos con jugadores humanos, véase Saijo y Nakamura (1995) y Rand y Nowak (2011).

... la neuroeconomía demuestra que el dinero no es la única divisa en el intercambio social. Recibir el regalo de un dólar es agradable, pero recibir la oferta de un dólar en el juego del ultimátum genera una reacción emocional negativa. (Hardy-Vallée 2007: 947).

En cuanto a los estados hedónicos a maximizar (como el placer, el bienestar, las emociones agradables, o la autoestima), tanto la microeconomía moderna como la neuroeconomía contemporánea están contribuyendo a un resurgir del sensualismo benthamita, que fue incorporado a la teoría económica moderna a través de revolución marginalista, y destacadamente con la influencia teórica de Edgeworth y Jevons. Ambos fueron buenos conocedores de las teorías psicológicas de su época, abiertamente decantadas por unos fundamentos hedonistas y consecuencialistas de la acción humana¹⁵⁸. Como explican Loewenstein, *et. al.* (2001), prácticamente todos los desarrollos de teoría de la elección (con o sin riesgo e incertidumbre) asumen una perspectiva cognitiva y consecuencialista, según la cual los agentes maximizan utilidad, y una de las expresiones de la utilidad es la expectativa emocional asociada a la decisión adoptada. En efecto, en microeconomía se han presentado modelos en los que se considera al agente racional como si fuera un productor-asignador idiosincrásico de sus propias emociones, con propósito maximizador (de emociones agradables), ya sea mediante la adquisición de bienes y servicios en el mercado, o a través de auto-provisión emocional. Las emociones entran en la función de utilidad como bienes de consumo (Frank 1988, Hirshleifer 1987; Becker 1997[1976b]), y el actor racional maximiza la experiencia hedónica de bienestar. En neuroeconomía también se han enfocado las investigaciones con premisas utilitarias similares (Sanfey, *et. al.* 2006; Seymour, *et. al.* 2007). De este modo, en la literatura económica la utilidad se especifica como variable dependiente en un algoritmo matemático descriptivo de la intensidad de una experiencia psicológica. No obstante, la cuantificación de la utilidad tiene valor a efectos ordinales, no cardinales. Sin embargo, aún no se ha descubierto el correlato empírico de la utilidad en el cerebro, a pesar de los intentos de identificar la utilidad con

¹⁵⁸ Aunque el hedonismo consecuencialista se instituyó con sus oportunos matices, sin embargo es considerado por Edgeworth una manera eficaz de aproximar la teoría económica del comportamiento humano a la metodología científica de las ciencias duras. De este modo, en *Mathematical Psychics* proclama explícitamente el principio del egoísmo hedonista como móvil conductual del comportamiento económico. La ventaja de este reduccionismo es aproximar metodológicamente la economía a la física: «*la concepción del Hombre como una máquina de placer* puede justificar y facilitar el empleo de términos mecánicos y de razonamiento matemático» (Edgeworth 2000[1881]: 56; cursivas en el original).

alguna modalidad científica de sensualismo psicológico, en este caso el asociado al sistema neuroendocrino de recompensa¹⁵⁹.

¿Es válida la teoría sensualista de la motivación? La neuropsicología ha demostrado que la motivación resulta activada fundamentalmente desde las emociones producidas en el sistema límbico, y que en la formación de las preferencias y la elección empírica el juicio deliberativo queda en segundo plano (Zajonc 1980). La amígdala controla el comportamiento mediante emociones actuando de forma incondicionada sobre el resto del cerebro, pero el resto del cerebro (el neocórtex, fundamentalmente, que es donde reside la mente lógica, la razón) no puede controlar de la misma manera a la amígdala, porque el circuito es distinto y evolutivamente más reciente (LeDoux 2000). Las emociones, pues, están omnipresentes en la vida humana, pero ¿actuamos motivados por emociones que conducen a experiencias de bienestar? Loewenstein y Lerner (2003) explican que el comportamiento humano está motivado por procesos psicológicos activados por emociones y que, en un momento dado, generan una disposición o necesidad de actuar de un determinado modo. Esta disposición es lo que motiva la acción, y no la expectativa de una experiencia emocional posterior. En el momento de la motivación, e incluso una vez iniciada la acción, pueden activarse nuevos procesos emocionales con sus correspondientes disposiciones motivacionales que, en ocasiones, entran en conflicto con las primeras. En algunos casos esto podría llegar a inhibir la acción iniciada. Al final del proceso, cuando se ha producido la acción, se activan los dispositivos neuroendocrinos de recompensa, que generan una experiencia emocional con valencia positiva o negativa en función del éxito o fracaso de la acción realizada.

La experiencia emocional positiva, que es una consecuencia de la acción, aunque no su causa, se debe a la producción y captación de neurotransmisores y neuropéptidos tales como la noradrenalina, la dopamina y la oxitocina (Sanfey, *et. al.*, 2006: 109-110). La “utilidad” con la que trabaja nuestro cerebro cuando refuerza un aprendizaje o un circuito innato según el grado de cumplimiento de objetivos se realiza estableciéndose

¹⁵⁹ Véanse los trabajos pioneros de Kahneman, *et. al.* (1997 y 1999), y otros más recientes como Helliwell y Barrington-Leigh (2010); Montage, *et. al.* (2006); Naqvi, *et. al.* (2006); Montage y Berns (2002).

una asociación entre comportamientos y experiencias emocionales pasadas. Así, se infiere una estrategia de comportamiento sobre la base de una memoria emocional en la que la magnitud de la recompensa puede experimentarse empíricamente en términos de bienestar, autoestima o emociones de significado positivo. Aunque todavía no se dispone de un conocimiento completo de las computaciones neurales subyacentes a este cálculo, existen investigaciones que implican al sistema de la dopamina (producida por las neuronas del área mesencefálica) en la señalización y la predicción de recompensas que conducen a motivar comportamientos a través del recuerdo de experiencias emocionales pasadas, y la adaptación y puesta al día de los objetivos y focos de atención en la memoria de corto plazo (Damasio, *et. al.* 1991; Naqvi, *et. al.* 2006). Se ha investigado igualmente el funcionamiento del sistema de recompensa en la motivación producida por lo que los economistas llaman utilidad “negativa”. Es el caso, por ejemplo, de la retractación de errores y la evitación de experiencias traumáticas, arriesgadas o simplemente desagradables (Tom 2007; Yeung y Sanfey 2004; y Holroyd y Coles 2002). Hay evidencia de que estos refuerzos negativos inhiben conductas y estimulan a buscar alternativas.

Esta operativa neurofisiológica que sí podemos llamar “sensualista”, sin embargo, *no equivale al hedonismo motivacional*, un aspecto éste último que puede considerarse un demostrado funcionamiento patológico del cerebro¹⁶⁰. En efecto, la búsqueda obsesiva de experiencias de bienestar conduce a refuerzos adictivos con suficiente poder como para inhibir otras emociones. Hace más de medio siglo que se demostró que las ratas de laboratorio a las que se implantan electrodos y se enseña la conducta de electroestimulación de los centros neurológicos del placer (los núcleos septales, o el haz prosencefálico medial) por ejemplo, a través de pulsar una palanca con su propio cuerpo, dejan de alimentarse y de beber líquidos, pierden el interés por la actividad sexual, y permanecen electroestimulándose hasta morir por inanición y estrés (Olds y Milner 1954). Comportamientos autodestructivos semejantes se producen tanto

¹⁶⁰ Las teorías actuales de la emoción se han desligado de las tradiciones filosóficas hedonistas, así como de las tradiciones que equiparaban las emociones a impulsos irrefrenables irracionales y perniciosos para la supervivencia del individuo. La psicología cognitiva contemporánea prefiere hablar de las emociones como experiencias psicológicas positivas (asociadas a la consecución de un objetivo) o negativas (cuando el objetivo no es logrado). Las emociones comparten con los instintos y con los hábitos el que son «disposiciones a actuar de una determinada manera», y esa acción es en sí misma una prioridad de ciertos objetivos o planes sobre otros objetivos o planes posibles. Véase Johnson-Laird y Oatley (1992).

en humanos como en animales con adicción a sustancias estupefacientes. Las ratas adictas a la electroestimulación, o los humanos toxicómanos, representan disfunciones patológicas del sistema de gestión hedónica del organismo, y por eso se rompe el equilibrio homeostático propiciando una dramática tendencia hacia la degradación. Los individuos sanos no pueden ser maximizadores de placer, dado que esa estrategia desemboca en la inhibición de la capacidad para sentir placer, y a medio y largo plazo se traduce en la incapacidad para la interacción social y la conducta apta para sobrevivir y reproducirse (Lüscher 2001; Bechara 2005). En otro orden de cosas, Fredrickson, *et. al.* (2013) han demostrado que la auto-gratificación hedónica persistente empeora el estado de salud, al estar interrelacionada con la sobreactivación del sistema del estrés.

El hedonismo motivacional fue también rebatido hace más de un siglo por William James, quien convino con Darwin y con Freud en la defensa de que existe una gran influencia de factores inconscientes e instintivos en el comportamiento humano. Para James, el asociar causalmente la acción humana a procesos de cálculo utilitario, en el sentido de balances hedónicos computados conscientemente y *a priori*, es una hipótesis disparatada: «¿Quién se ríe por el placer de reír, o frunce el ceño por el gusto de hacerlo? ¿Quién se sonroja para evitar la incomodidad de no sonrojarse» [James (1920[1892]), citado en Stigler (1979[1965]: 81). James está enfatizando que la teoría hedonista de la motivación está amparada en una causalidad espuria. En efecto, como señala Wispé (1991: 81) metafóricamente: «Volar en avión siempre viene acompañado de consumo de combustible, pero no volamos en avión para consumir combustible». Adam Smith objetó una observación similar en *La teoría de los sentimientos morales*:

... tanto el placer como el dolor son experimentados siempre de forma tan instantánea, y a menudo bajo circunstancias tan frívolas, que parece evidente que no puede derivarse de tales consideraciones el propio interés. (Smith 1997[1759]: I, i, 2, p. 57).

La psicología evolutiva no admite la hipótesis de que los animales, y en particular los humanos, actúen motivados por la consecución de beneficios hedónicos, ni que maximicen premeditadamente ninguna variable utilitaria a través de la lógica y el cálculo racional. Los seres vivos “toman decisiones” que repercuten positiva o negativamente en su eficacia biológica. A lo largo del tiempo los procesos de evolución seleccionan los patrones decisorios que proveen de mayor eficacia. Esos patrones

decisorios acaban implantándose filogenética y/u ontogenéticamente en dispositivos neurológicos que se activan ante situaciones similares que requieren resolver problemas concretos de supervivencia. No obstante, tales disposiciones conductuales no necesariamente son siempre las óptimas en cada circunstancia. Los individuos toman decisiones sin saber de antemano si son las mejores, y lo hacen con una gama amplia, aunque limitada, de programas neurológicos entre los que puede figurar el más apropiado, o también pueden ser inapropiados para el problema específico surgido de condiciones ambientales para las que no existe programación implantada (Gintis 2003; Campbell 1972 y 1974). En buena lógica darwinista, los programas neurológicos adaptativos tienden a replicarse con éxito, mientras que la propagación de los no adaptativos tiende a ser imposibilitada por los procesos evolutivos (genéticos y culturales). De este modo, los organismos no son agentes razonadores prácticos que se propongan maximizar utilidad o eficacia biológica, sino *ejecutores de programas cognitivos* generados en el pasado de la especie y/o del organismo, que resultaron adaptativos para los individuos que los transmitieron por herencia genética y/o cultural (Cosmides y Tooby 1994a, y 2001).

En la teoría evolutiva de la racionalidad, entendida ésta como eficacia en la supervivencia y reproducción, la elección racional es un *a posteriori* del comportamiento animal, y no un *a priori* relacionado con la elección de objetivos y procedimientos de manera consciente, intencional y deliberada (Simon 1983: 37)¹⁶¹. La acción de los organismos no se produce ni completamente a ciegas ni tampoco con una racionalidad optimizadora, sino con regularidades de eficacia probada, aunque siempre con un determinado nivel de incertidumbre. Esos programas cognitivos son patrones de acción fijos (*hard wired mechanisms*) o aprendizajes culturales (*imprintings*), y la única diferencia que existe entre los humanos y el resto de animales es que los primeros disponen de un cerebro capaz de albergar más instintos porque disponen de una mayor cantidad de sinapsis neuronales, y por lo tanto una mayor complejidad y plasticidad:

¹⁶¹ Así, los individuos que tienen éxito en la supervivencia y reproducción son aquellos que las condiciones ambientales han permitido que lo tengan, y no tanto los que han “tomado las decisiones” adecuadas para sobrevivir en ese medio. El proceso que ha conducido a que un determinado rasgo haya sobrevivido ha tenido lugar porque los antepasados de los individuos actuales fueron capaces de sobrevivir «disparando (actuando) a bulto», como si lo hicieran con una «escopeta mal calibrada», pero sus disparos dieron en los blancos adecuados para sobrevivir (expresiones tomadas de Alchian 1950: 219). Véase Colman (2005).

«Nuestra alabada flexibilidad proviene de multitud de instintos ensamblados en programas y que entre sí mantienen competiciones» (Pinker 2004: 244). Esos programas pueden resultar útiles para conseguir objetivos instrumentales; sin embargo, parafraseando nuevamente a William James:

La ciencia puede muy bien llegar y examinar estas vías, y hallar que casi todas son útiles, pero las criaturas no las siguen por su utilidad, sino porque en el momento en que las siguen sienten que hacen precisamente lo único que es apropiado y natural hacer. [James (1920[1892]), citado en Pinker (2004: 245)].

2.5.- Conclusiones

Se han podido dilucidar aspectos específicos de la naturaleza del sujeto y de los procesos psíquicos que surgen del organismo y constituyen la mente. La principal conclusión a la que hemos llegado, una vez examinada la literatura neurocientífica y contrastadas las teorías filosóficas más significadas, es que, como indican Castro Nogueira, *et. al.* (2003: 184):

El enfoque evolutivo de la mente muestra, sin mucho lugar para las dudas, que la concepción del ser humano característica de las ciencias sociales se ha construido sobre presupuestos incompletos.

En efecto, el yo (*self*) teorizado por la filosofía hasta no hace mucho tiempo, y aceptado como axioma por una parte de las humanidades y ciencias sociales, consiste en un epifenómeno ilusorio. No existe un sujeto razonador que decida racionalmente y de forma autónoma lo mejor (lo racional) para sus propios intereses racionales:

Se trata de un concepto, una construcción de la razón y de la imaginación, una manera de ver y componer el mundo, pero no una entidad que exista realmente. (Kolm 1987: 257).

El yo, por lo tanto, es un producto de la mente, y no ésta un producto de la actividad del yo. Las sensaciones, sentimientos, ideas, imaginaciones, emociones y voliciones *son* la propia mente, y no atributos de un sujeto que los maneja y domina (Kolm 1987: 257).

Pocos análisis filosóficos del pasado atisbaron estas conclusiones y las situaron en unas coordenadas teóricas aprovechables para la ciencia. Algunas de estas contadas excepciones son los casos, por ejemplo, de Spinoza y Hume. Ambos autores sí

dedujeron la naturaleza del sujeto como un producto de la mente que genera una ilusión, aunque sólo Spinoza aportó una teoría de cómo el funcionamiento orgánico puede, a través de los instintos, producir ese espejismo. Tanto Spinoza como Hayek disponen de una teoría evolutiva de la mente, y la presentan como un compendio de procesos orientados a lo que hoy día diríamos eficacia (biológica y cultural). Sus aportaciones han sido reconocidas por la neurología contemporánea, y suponen hitos históricos en la comprensión del cerebro y sus funciones vitales¹⁶².

La subjetivación, el individuo consciente, consiste pues en un epifenómeno neurobiológico consecuencia de la integración de procesos neurales de alto nivel. Actualmente es posible explicar el problema fuerte de la biología sin necesidad de hipótesis *ad hoc* ni prejuicios animistas. Lo mismo se puede decir del origen de la motivación humana y de los procesos decisorios que tienen lugar en el cerebro:

Usar la racionalidad como una explicación de la elección es similar a usar el animismo como explicación de la capacidad de los animales para moverse -(una teoría) admirablemente descriptiva, pero que falla a la hora de seguir la pista o conectar con la ciencia de la causación real. (Tooby, *et. al.* 2006: 126n₁; paréntesis añadido).

La ciencia ha llevado a cabo descubrimientos que matizan o derriban teorías bien asentadas desde hace siglos acerca del psiquismo humano y de su influencia en el comportamiento. Freud, por ejemplo, rechazó de plano en *La interpretación de los sueños* (1900) la posibilidad de que alguna vez se llegase a identificar estructuras anatómicas donde se localizan procesos mentales, o que los procesos mentales se correspondieran con estructuras físicas específicas (Fischer 1990: 7). Sin embargo, un siglo después, sabemos que la opinión de Freud ha sido refutada y superada. Del mismo modo, hoy también podemos asegurar que el yo, entendido como unidad autónoma de conciencia, conocimiento, deliberación racional, evaluación y decisión, es un mito respaldado en un espejismo, aunque con una gran capacidad de influir en las creencias de la persona o individuo social; esto explica su carácter universal. Sin embargo, aunque ilusorio, el yo es una función adaptativa, como lo demuestra el que los individuos en estados permanentes de conciencia sin yo, o con yoes patológicos, tienen una eficacia biológica drásticamente disminuida.

¹⁶² Véanse Damasio (2005) sobre la obra de Spinoza, así como el prólogo de Joaquín M. Fuster a la edición que aquí utilizamos de *El orden sensorial* en Hayek (2004[1952]).

Como implicación adicional, se puede señalar que algunos de los problemas filosóficos clásicos de la epistemología, como por ejemplo la validación de la inducción, el solipsismo y las otras mentes, o la fiabilidad de las percepciones sensoriales; además de los relacionados con la filosofía moral y política relativos a la dimensión prosocial o pro-individual, cooperativa o competitiva, egoísta o altruista de la naturaleza humana; y también algunas de las anomalías y controversias que ocupan los debates teóricos dentro de las ciencias sociales, como las relativas a la racionalidad, los sesgos cognitivos y decisorios, las reglas de conducta en situaciones de riesgo o de incertidumbre, el papel jugado por las emociones en el comportamiento, etc., tienen su explicación en el funcionamiento del órgano que produce la cognición y los comportamientos, que es el cerebro.

El cerebro, como cualquier otro órgano, y como el organismo humano en su conjunto, es un producto evolutivo adaptado al ambiente en el que se desarrollaron los antecesores pre-humanos y humanos en la sabana africana hace decenas de miles de años (Kanazawa 2004a). Concretamente, el cerebro de la especie *Homo sapiens* es un diseño neurofisiológico de la evolución que ha permanecido sin modificaciones sustanciales durante los últimos cuarenta o cincuenta mil años. Todos los dispositivos orgánicos y neurológicos (como el cerebro, el intestino o el ojo), y los resultados de su funcionamiento, tienen una racionalidad adaptativa, y por tanto ecológica. Se trata de apriorismos de la naturaleza humana. Estos dispositivos orgánicos existen porque son adaptaciones eficaces al medio en que el evolucionaron. Por tanto favorecen la supervivencia y la replicación de las unidades replicadoras que portan los organismos adaptados a ese medio (Barkow, *et. al.* 1992).

¿Qué sentido innato tiene pues el comportamiento animal? ¿Hay alguna diferencia entre el comportamiento humano y no humano? Durante los dos últimos milenios se ha planteado que el comportamiento humano, a diferencia de los demás seres vivos, es un comportamiento racional. La racionalidad, un concepto controvertido, ha sido presentada como deliberación lógico-lingüística, cálculo, coherencia, egoísmo e intencionalidad consecuencialista (y optimizadora) de la acción. Todas esas acepciones, imputadas tradicionalmente sólo al ser humano, están siendo matizadas o cuestionadas

por las neurociencias. Tanto los humanos como los demás seres vivos están equipados para llevar a cabo procesos de optimización, aunque esa capacidad sea siempre limitada y se ejecute inconscientemente (Gintis 2007b: 191). La racionalidad no es una habilidad del yo, sino un conjunto de procesos orgánicos funcionales presentes en todos los seres vivos. Estos procesos operan con fluidez y armonía y sin sensación de esfuerzo alguno para la gestión de la información que afecta al organismo, y lo hacen de tal manera que su existencia y operatividad pasan desapercibidas a la conciencia. Los procesos neurales son variados y especializados, y entre ellos caben destacar: razonamiento sobre objetos, causalidad física, cálculo, inferencia sobre creencias y motivaciones de otros individuos, e interacciones sociales (Thornhill, *et. al.* 1997: 227). Todo ello avala la hipótesis de una racionalidad instintiva universal, integrada por una amplia y compleja diversidad de instintos racionales armonizados (Pinker 2004: 389 y ss.).

El carácter racional de un comportamiento no se basa pues en una intencionalidad racionalizadora, que contempla elecciones y decisiones conscientes sobre fines y medios. La racionalidad instintiva no es una fuerza que empuja al comportamiento, sino una que tira de él, un atractor pre-establecido en el diseño neurológico adaptativo:

Tanto los humanos como los demás animales razonan, deciden y se comportan en virtud de ingenios computacionales incorporados en el sistema nervioso. Por ello, cualquier explicación causal del comportamiento que sea -racional o de otro tipo- necesariamente invoca a teorías acerca de la arquitectura de esos ingenios computacionales. La racionalidad de un comportamiento resulta irrelevante a efectos de su causa o explicación. (Cosmides y Tooby 1994b: 327).

Es preciso comprender además el papel de las emociones en el organismo, y su intervención en los procesos racionales. Las emociones, como los demás programas innatos y aprendizajes culturales implantados en la ontogénesis (*demonos*), están orientados a la homeostasis del organismo, y la causa última de la homeostasis es evitar la degradación orgánica sin replicación. En este sentido, y desde una perspectiva evolutiva, la racionalidad es un producto adaptativo al ecosistema de la especie, por lo tanto un producto ecológico, como explica Vernon Smith:

Hay un sentido en el que los sistemas ecológicos, ya sean culturales o biológicos, deben ser necesariamente, o están en el proceso de ser, racionales: sirven a las necesidades de eficacia de aquellos que los crearon de manera no intencional a través de sus interacciones. (Smith 2002: 511).

Si el comportamiento es racional en el sentido de ecológicamente adaptativo, y un comportamiento bien adaptado es aquél que permite al organismo replicarse con éxito, entonces, si el sujeto es una ilusión, ¿cuál es el interés que persiguen las decisiones de los individuos? ¿A qué o quién van orientadas? No tiene sentido científico identificar la racionalidad con el egoísmo psicológico que elige maximización de pagos pecuniarios o hedónicos. Tampoco se puede afirmar que los organismos actúen motivados por un sensualismo consecuencialista:

La distinción formal básica entre la racionalidad económica y la evolucionista se basa en que la primera explica los patrones de comportamiento como consecuencias de la optimización *ex-ante* de los agentes, mientras que la segunda explica los patrones de comportamiento como consecuencias *ex-post-facto* de comportamientos ya realizados, que predeterminan el comportamiento subsecuente dentro de juegos de la misma estructura. (Herrmann-Pillath 1994: 48).

Consecuentemente, la racionalidad ecológica, aunque parece brotar de las entrañas sin ninguna operación mental voluntaria, es sin embargo un fenómeno mucho más complejo que el esbozado en las interpretaciones reduccionistas. Las emociones y programas cognitivos implantados intervienen en el propósito ecológico-racional de preservar la vida del individuo en orden a maximizar su reproducción, siempre que, claro está, tales adaptaciones estén bien ajustadas al medio en el que el individuo se esté desarrollando.

El medio de referencia en el que se desenvuelve *Homo sapiens* es un medio social, que es además nicho cognitivo y cultural. La adaptación del yo a ese medio a través de la interacción social redescubre al sujeto en una dimensión relacional que pone en tela de juicio las tradiciones filosóficas solipsistas e individualistas, tal y como explican Ramachandran y Blakeslee (1999: 317),

... a la sociedad (por lo general, parientes que comparten nuestros genes) le resulta más fácil incorporarnos a sus planes, aumentando así nuestras posibilidades de supervivencia y perpetuación de nuestros genes (...) Y aquí tenemos la mayor ironía: que el yo, que casi por

definición es algo completamente privado, es en gran medida un artefacto social, un cuento que inventamos para los demás.

Curiosamente, el comportamiento prosocial humano tiene como pre-condición psicológica la aparición de la conciencia, un evento natural consistente en el arraigo filogenético de una poderosa creencia en que existe un observador que guía el cuerpo y se relaciona con su entorno de forma inteligente. Dentro de ese entorno existen otros seres, que reconoce y comprende, con los que incluso se identifica, aunque la experiencia de autonomía individual siempre está presente (Hoffman 2000: 71). La aparición del ego, el supuesto epítome del subjetivismo y del egoísmo, es, paradójicamente, aquello que posibilita la faceta inherentemente humana del altruismo moral y psicológico realizado con plena consciencia e intención.

Capítulo 3.- Alcance y límites de las explicaciones basadas en los conceptos de cooperación recíproca y egoísmo racional

Necesitamos una teoría evolucionista de la moral (...) (Su principal aportación será que la moral no es una creación de la razón, sino una segunda tradición independiente de la tradición de la razón, que nos ayuda a solucionar problemas que exceden en mucho los límites de nuestra capacidad de raciocinio. (Friedrich von Hayek, en un discurso pronunciado en Oesterreichisches College, Viena, 1985. Citado en Zak 2008a: 259).

Los comerciantes que viajan por mar siguen una conducta prudente cuando procuran que los que van en el mismo barco no sean disolutos, blasfemos ni malos. (Montaigne, *Ensayos*, Libro I, Capítulo XXXVIII, “De la soledad”).

No tiene ningún mérito portarse bien con otra persona. Teresa tiene que ser amable con los demás aldeanos porque de otro modo no podría vivir en la aldea. Y hasta con Tomás *tiene* que comportarse amorosamente, porque a Tomás lo necesita. Nunca seremos capaces de establecer con seguridad en qué medida nuestras relaciones con los demás son producto de nuestros sentimientos, de nuestro amor, de nuestro desamor, bondad o maldad, y hasta qué punto son el resultado de la relación de fuerzas existente entre ellos y nosotros. (Kundera 2004: 295; cursiva en el original).

Tengo el presentimiento de que llegará el momento en que consideremos la invención del concepto de estrategia evolutivamente estable como uno de los avances más importantes en la teoría evolutiva desde Darwin. (Dawkins 1993[1976]: 110; paréntesis añadidos).

3.1.- Introducción

El altruismo es un hecho empírico constatado, que ha sido analizado desde diversos enfoques teóricos, y sobre el que sólo hace relativamente poco tiempo se han formulado explicaciones de tipo científico. En el presente capítulo se presentan los análisis del altruismo más tradicionales y divulgados durante las últimas décadas, y se explora tanto su alcance, como las fortalezas y principales debilidades de estos estudios.

En su expresión más básica el altruismo humano puede definirse como la acción consciente, y con conocimiento de sus repercusiones en sí mismo y otro/s, que consiste en beneficiar a otro/s sin la intención de obtener por ello alguna ventaja personal o eludir algún castigo, ya sea directa o indirectamente (De Quervain, *et. al.* 2004). El altruismo conlleva un compromiso conductual con un patrón de comportamiento cooperativo incluso en casos en que se puede obtener ventaja personal anónimamente, o sea, sin que exista posibilidad de castigo (Frank, *et. al.* 1993b: 249). Esta acepción incluye la conducta de no aprovecharse de la vulnerabilidad ajena cuando ello supone la renuncia a conseguir alguna ganancia (Stout 2008: 162). Tales manifestaciones de altruismo pasivo, o basal, son compatibles con el espectro de acciones clasificadas entre lo que Simon (1983: 57-58) denomina altruismo *fuerte* (que conlleva un sacrificio de eficacia biológica para quien lo lleva a cabo) y altruismo *débil* (que consiste en un sacrificio a corto plazo no condicionado a reciprocidad, si bien con una futura aunque incierta posibilidad de ganancia)¹⁶³. Cuanto más cercanas estén las consecuencias de una acción al pago de un coste medido en eficacia individual, para que otro/s individuos obtengan un beneficio, más altruista en sentido fuerte será dicho comportamiento¹⁶⁴. Esta conducta puede ser considerada una anomalía evolutiva, y plantea un problema teórico.

¹⁶³ Frank, *et. al.* (1993: 249) se refieren al mismo fenómeno explicitado por Simon, utilizando la terminología de *hardcore altruism* o *hardcore cooperation*, para referirse a la acción de cooperar en el dilema del prisionero de única jugada, siendo *softcore cooperators* aquéllos que sólo cooperan en dilemas del prisionero repetidos a futuro. Por el contrario, los defectores son puros oportunistas que nunca cooperarán en el dilema del prisionero si de ello pueden aprovechar ganancia a corto plazo.

¹⁶⁴ Téngase en cuenta que los dos polos opuestos de la distribución de los comportamientos en las relaciones entre organismos son el altruismo (fuerte) y el conflicto, entre los cuales se sitúan diferentes grados de cooperación (Tullock 1990: 158). Aquí, sin embargo, vamos a clasificar las acciones en dos grupos, altruismo y egoísmo, dentro de los cuales cabe igualmente una graduación. El ejemplo más obvio de altruismo consiste en perder recursos en favor de otro sin obtener ganancia o eludir algún castigo por hacerlo. El ejemplo típico de egoísmo extremo es detraer recursos de otro en un juego de suma negativa. El conflicto se plantea siempre que existe un juego de suma cero, o negativa, y ambos agentes son egoístas.

Desde la perspectiva fuerte el altruista realiza externalidades positivas sobre otros sin internalizarlas. El altruismo débil, por el contrario, puede consistir en beneficiar a otros sin pagar costes personales ni obtener beneficios directos e inmediatos. Dado que no se obtiene una ganancia neta y cierta a corto plazo, el altruismo (fuerte y débil) tiene una motivación intrínseca, no-consecuencialista y no-egoísta. El altruista genuino actúa por la motivación misma de hacerlo (Andreoni 1993). Una consecuencia del comportamiento altruista puede ser recibir *a posteriori* una gratificación psicológica que no es la causa próxima de la acción, sino una consecuencia impremeditada (Dunn, *et. al.* 2008; Harbaugh, *et. al.* 2007; Moll, *et. al.* 2006).

Al contrario del altruista, un agente intencionalmente egoísta sí internaliza ventajosamente las externalidades positivas que realiza, o evita si puede dicha internalización en el caso de las externalidades negativas que pueda ocasionar. Esto es así porque en la conducta egoísta el individuo maximiza su eficacia sólo en cuanto a individuo que es, como si no formara parte de una población dotada de procesos regulatorios de la supervivencia. La motivación que impulsa su comportamiento consiste exclusivamente en la cuantía de sus propios beneficios individuales. Dado que el egoísta no considera los costes e ingresos ajenos, su proceder está referido sólo a sí mismo, a su propia mente y sus propios deseos. El egoísmo puede evolucionar porque, *ceteris paribus*, al estar asociado a la expectativa de un beneficio cierto, confiere eficacia diferencial respecto al altruismo comprometido, en el que no existe beneficio, sino coste (altruismo fuerte), o bien un beneficio incierto (altruismo débil).

¿Cómo puede haber evolucionado el altruismo si en interacción con el egoísmo produce aparentemente una replicación inferior en quien lo realiza? Planteado de otro modo: ¿cómo puede explicarse teóricamente una conducta que evolutivamente juega en contra de quienes así se comportan? Esta pregunta es uno de los interrogantes científicos de la teoría evolutiva, y concretamente un problema teórico desde el punto de vista de la psicología evolutiva y la neurobiología (Tomasello, *et. al.* 2012: 673). Ello se debe a que las motivaciones intrínsecas que impulsan la generosidad altruista no tienen razón de ser, como causas próximas, si no están respaldadas por una causa última determinante que aporte eficacia diferencial.

El problema teórico ya fue expresado por Darwin en *El origen del hombre*. Se trata de la paradoja evolutiva de los comportamientos no-exclusivamente-egoístas, que incluyen la gama de conductas que delimitan en el extremo con el exclusivo interés propio, pasando por diferentes niveles de cooperación y generosidad, hasta llegar al altruismo heroico a favor de una causa que no repercute favorablemente en la eficacia del altruista o su progenie, y que incluso puede perjudicar. La paradoja, según la formuló Darwin, es la siguiente:

Cuando dos tribus de hombres primitivos que viviesen en el mismo país compitiesen entre sí, dado que en otras cualidades seguramente habría paridad, la victoria estaría de parte de aquella compuesta por individuos valerosos, más simpáticos y fieles, dispuestos siempre a avisarse mutuamente de los peligros y a defenderse y ayudarse (...) (Pero) es sumamente dudoso de los padres más simpáticos y benévolos, o de aquellos que fuesen más fieles para con sus compañeros, haya resultado mayor descendencia que de los padres egoístas y falsos, pertenecientes a la misma tribu. Si el individuo generoso estuvo siempre dispuesto a sacrificar su vida antes que hacer traición a sus camaradas, es fácil que pierda en la demanda de la vida sin dejar herederos de su noble conducta y naturaleza (...) Por lo cual, parece casi imposible que el número de hombres dotados de las sobredichas virtudes, o el grado de estas mismas virtudes, hayan aumentado por selección natural, esto es, por sobrevivir los más aptos. (Darwin 1972[1871]: I, v. p. 123, y 125-126; paréntesis añadido).

Lo que está resaltando Darwin es el problema clásico de la acción colectiva, tal y como específica Price (2008: 230): en una interacción cooperativa que implica producción conjunta, los que no cooperan obtendrán, *cæteris paribus*, una superior ventaja evolutiva derivada de las externalidades positivas creadas por los cooperadores en la dinámica grupal. O dicho de otro modo: «Dado que la acción colectiva beneficia a todos en el grupo, contribuyan o no, la selección natural favorece a los que no contribuyen» (Boyd y Mathew 2007: 1.858). Aunque en el largo plazo cooperar puede resultar rentable (si existen cooperadores con quienes poder cooperar en interacciones futuras), sin embargo en el corto plazo quien actúa como un parásito o un depredador sacaría ventaja antes de que el cooperador llegue a obtener sus inciertos beneficios futuros. Se trata de una versión de la mayoría de juegos que representan dilemas sociales, como el del prisionero, el de la confianza o el de la provisión de bienes públicos. En este sentido, el comportamiento de quien se expone a ser explotado o depredado a corto plazo por otros es evolutivamente disfuncional.

La pregunta que se deriva de la paradoja del altruismo consiste pues en ¿por qué los individuos contribuyen en los dilemas sociales si no haciéndolo, y aprovechándose

de los que sí contribuyen, pueden mejorar comparativamente su eficacia? Si los cooperadores no pueden atajar este problema de pérdida relativa de eficacia, tenderán a ser explotados y se extinguirán. ¿Qué soluciones han planteado las ciencias del comportamiento a esta encrucijada?

Dos soluciones teóricas tradicionales dadas a este problema evolutivo se basan, en primer lugar, en negar el altruismo y afirmar el supuesto de un egoísmo inteligente universal, o propio interés racional a largo plazo, que consiste en un egoísmo prudente, premeditadamente restringido a fin de obtener beneficios estratégicos o evitar males personales mayores (Gauthier 1994[1986]; Mises 2005[1944]). Los individuos que no saben ser egoístas, de este modo, es que o bien no comprenden el significado del juego o es que no han aprendido a jugar: necesitan más experiencia para aprender a jugar la estrategia dominante en dilemas sociales (Andreoni y Croson 2008: 776). En segundo lugar, la otra solución teórica ampliamente utilizada es admitir una auténtica disposición altruista si bien sujeta a una férrea reciprocidad, o sea, una cooperación condicional supeditada a que el receptor del altruismo se comporte también altruistamente, ya sea en el presente o en el futuro (Fischbacher, *et. al.* 2001; Back y Flache 2006).

Hirshleifer (1978: 242) ha planteado la existencia de cinco grupos de soluciones evolucionadas –que aquí se presentan ordenadas en cuatro– para atajar el problema del oportunismo en interacciones llevadas a cabo entre cooperadores racionales. Estos últimos son individuos que actúan en última instancia motivados por objetos instrumentales al servicio de su propia eficacia. Tal supuesto conlleva que los deseos remotos que motivan la acción están siempre orientados a su eficacia personal. Los cooperadores instrumentales cooperan por conveniencia, al estilo figurado del *homo oeconomicus*, y en último término dicha conveniencia resulta ventajosa, aunque en casos no aporte más que una gratificación psicológica de utilidad¹⁶⁵. Dichas soluciones favorecen la evolución de la cooperación entre organismos que maximizan su eficacia sujeta a restricciones. Se trata de las siguientes:

¹⁶⁵ Lógicamente, la hipótesis de la gratificación psicológica como causa próxima de la motivación es tautológica si no viene explicada por una causación última con sentido evolutivo. Por otra parte, aun admitiendo la veracidad de la experiencia hedónica asociada a la cooperación, suele confundirse la causa con la consecuencia de ésta: «la intuición correspondiente es que, puesto que soy yo quien actúa, incluso cuando actúo en interés de otro, debe haber un interés mío que proporcione el impulso» (Nagel 2004[1970]: 90).

- (1) Cooperación exclusivamente con parientes genealógicos, que nos sitúa en la regla de Hamilton y la estrategia optimizadora de la población de réplicas genéticas.

- (2) La cooperación condicional fuera del ámbito familiar, o cooperación condicional entre individuos optimizadores de su eficacia individual que practican la regla de reciprocidad¹⁶⁶. Se trata de una relación simbiótica entre dos o más individuos, independientemente de que estén emparentados o no y de que pertenezcan a la misma o a distinta especie. Esta solución adaptativa aporta incentivos a cooperar entre egoístas que actúan como si seleccionaran una estrategia eficaz para producir equilibrios cooperativos. Dentro de ella se apuntan dos posibles marcos relacionales:
 - (2.1) Compromiso condicional: consiste en el modelo de la reciprocidad directa del tipo toma y daca (*tit for tat*) (Axelrod 1996) en el que la amenaza creíble de la respuesta de reciprocidad (positiva o negativa) disuade la tentación (propia o ajena) de comportarse como un aprovechado. Aunque es racional cooperar de este modo, este patrón de cooperación requiere adaptaciones psicológicas de refuerzo tales como emociones viscerales (p.ej. la indignación, la ira y el odio), que en un planteamiento ultrarracional y estático serían consideradas irracionales (Frank 2001). Paradójicamente, este mecanismo próximo de refuerzo de la cooperación, aparentemente anti-cooperativo, permitiría explicar la supervivencia de un rasgo (el de la agresión) aparentemente disfuncional para la cooperación.

¹⁶⁶ David Gauthier se refiere con el término de maximización restringida al principio lockeano de restricción de la libertad de acción en la persecución regulada del propio interés:

Distinguimos entre la persona que está dispuesta a maximizar directamente su satisfacción o a lograr su interés en las elecciones particulares que haga, y aquella otra persona que está dispuesta a acatar las restricciones morales mutuamente ventajosas, con la condición de que espere un acatamiento similar por parte de los demás. Esta última persona es lo que llamamos un maximizador restringido. Y los maximizadores restringidos, al mantener una interacción con los demás maximizadores restringidos, gozan de oportunidades de cooperación de las que carecen aquellos otros (los que no someten a ninguna restricción su propio interés). (Gauthier 1994[1986]: 32; paréntesis añadido).

- (2.2) Cooperación social intra-démica: actuar sólo dentro de estructuras sociales o marcos institucionales que favorecen la aparición de interacciones repetidas, y que por lo tanto permiten el ejercicio de la reciprocidad directa y la toma en consideración de las reputaciones (reciprocidad indirecta). Así, cuando la reputación que identifica al individuo tiene valor informativo, los resultados futuros de no cooperar, descontados al presente, desincentivan a actuar de ese modo.
- (3) Interacción basada en la interdependencia mutualista: «según el grado con que los grupos comparten un destino común, ayudar a los otros es ayudarse a sí mismo» (Hirshleifer 1978: 242). Un ejemplo que respalda esta teoría del destino común es el de los animales de naturaleza grupal, como la abeja, el delfín o la especie humana. Si los individuos rehúsan ayudar a los otros miembros del grupo pueden llegar a extinguirse, porque dependen unos de otros para sobrevivir y reproducirse. Cuando los grupos productivos humanos se amplían pueden aprovechar las ventajas de una fuerte interdependencia recíproca, tanto de complementariedad como de suplementariedad. La clave evolutiva que explica el mecanismo próximo de la vinculación, en este caso, son las sinergias derivadas de la especialización y la cooperación cuando se trata de unidades grupales con una alta especificidad de activos y conocimiento (primer caso), o economías de escala (en el segundo caso).
- (4) La identidad cultural, o sea, el comportamiento ético y el adoctrinamiento cultural. Según Hirshleifer tal vez sólo en la especie humana haya evolucionado esta capacidad, consistente en la aceptación conformista de la influencia social a través de la transmisión cultural de valores, creencias y estereotipos de conducta en la ontogénesis. Estos replicadores constituyen las identidades biológicas o “tribales” con que los individuos se reconocen a sí mismos, y en la especie humana se da un predominio cultural en la conformación y propagación de replicadores que explica las identidades exogámicas. La cultura se convierte en un factor de compromiso con la

cooperación no por la motivación del propio interés, sino por las capacidades innatas de identificación cultural y conformismo acrítico características de la especie. Además, dicha transmisión de pautas cooperativas de comportamiento genera dinámicas sociales de evolución cultural activadoras de procesos de selección grupal.

Las soluciones (1), (3) y (4) pueden considerarse ejemplos de altruismo, puesto que se requieren adaptaciones filogenéticas pre-rationales que constriñen por sí mismas la posibilidad de un único egoísmo estratégico como causa motivadora de la acción (Tomasello y Vaish 2013; Tomasello, *et. al.* 2012). Sin embargo, aunque la solución (2), en sus dos variantes, es igualmente el producto de adaptaciones pre-rationales con un sentido de maximización restringida, y no sólo individualista, aquí vamos a tratarla como el caso más próximo a la noción de egoísmo impuro, coherente con el supuesto del egoísmo racional y deliberadamente condicionado (Ostrom 2000: 140). En el presente capítulo se analiza el conjunto de soluciones del tipo (2), y se lleva a cabo una valoración crítica de la posibilidad de una única explicación evolutiva de la cooperación basada en el interés propio a largo plazo, propio interés prudencial, o egoísmo racional¹⁶⁷.

Si bien en la microeconomía de la no-cooperación se considera que todos los individuos son oportunistas potenciales, en el presente capítulo se introduce una taxonomía consistente en que dentro de la población existen dos tipos de conducta: por un lado individuos que cooperan con otros individuos aplicando la regla de reciprocidad (cooperadores condicionales); y por otro lado no-cooperadores (defectores, *free-riders* y oportunistas en general) que tienden a no cooperar, y que cuando lo hacen es autolimitando su propio interés por conveniencia si los incentivos hacen rentable esta estrategia (Morishima, *et. al.* 2012)¹⁶⁸.

¹⁶⁷ Entiéndase por egoísmo racional la conducta de perseguir los propios objetivos con perspectiva a largo plazo y de una manera consistente, sean cuales sean esos objetivos. El agente racional maximiza su resultado *individual* esperado, computando el descuento del futuro (Binmore 2006a: 8).

¹⁶⁸ La cautela metodológica del oportunismo potencial tiene un antecedente filosófico en el que Hausman (1998: 67) llama “axioma de Hume”, quien en un ensayo político publicado en 1741 afirmó que:

Los escritores políticos han establecido como máxima que en la concepción de cualquier sistema de gobierno, y en la fijación de los diversos sistemas de equilibrios y controles de la Constitución, cada hombre debe ser considerado un truhán, que no tiene otro objetivo, en todas sus acciones, que el interés privado. (Citado en Hausman 1998: 67).

Más de dos siglos y medio después de Hume, su prescripción sigue viva en la fundamentación teórica

Dentro de este contexto, vamos a explicar cómo en ecosistemas sociales en los que existen cooperadores y no-cooperadores la cooperación sólo puede sobrevivir si es condicional y se apoya en la regla de reciprocidad, y además existe una elevada probabilidad de que las interacciones entre los individuos se vuelvan a repetir (apartados 2 y 3). Expondremos sintéticamente los principales modelos de reciprocidad formulados en la literatura especializada, con especial dedicación a la reciprocidad directa (sub-apartado 3.1) y la reciprocidad indirecta (sub-apartado 3.2).

La primera es el arquetipo de los intercambios diádicos en juegos de suma positiva, y la segunda es un refuerzo adicional de la cooperación en los sistemas sociales (una especie de antídoto institucional contra el oportunismo) tras la cual han evolucionado los sistemas de reputación, y las regulaciones morales y legales, así como la institución del dinero. Se explica cómo la reciprocidad indirecta requiere que existan cooperadores capaces de castigar con una represalia a los no-cooperadores, aunque esa externalidad positiva no la internalicen, e incluso aunque no hayan sido perjudicados directamente por éstos.

El capítulo concluye con un análisis dinámico de los problemas asociados a los modelos de cooperación basada en la reciprocidad y el egoísmo racional (apartado 4). Para que estos modelos sean empíricamente explicativos se precisa un ambiente y estructura poblacional estables, de pequeña dimensión, es decir, un agregado social en el que los individuos puedan volver a encontrarse en interacciones futuras, y que la información sobre las reputaciones sea abundante, descentralizada y fiable. La cuestión clave en este sentido consiste en si es suficiente con los patrones conductuales de reciprocidad y egoísmo estratégico para que la cooperación pueda evolucionar en sistemas sociales complejos y no estacionarios. Pasemos a analizar la capacidad explicativa de este tipo de modelos en la actualidad.

de la microeconomía, como en Williamson (1993: 485): «El mundo del comercio se reorganiza a favor de los cínicos y en contra de los inocentes cuando los científicos sociales emplean un lenguaje amistoso que no es descriptivamente exacto». Gary Becker (1987: 233-234), por su parte, explica con su teorema del niño malo (*rotten kid theorem*) que incluso puede darse el caso de egoístas estratégicos que persiguen su propio interés simulando ser altruistas.

3.2.- Cooperación dinámica entre egoístas racionales: del Leviatán filantrópico a la reciprocidad

El oportunismo y la defección en los dilemas sociales son un problema clásico en el análisis de la acción colectiva. En juegos como el dilema del prisionero y la provisión y sostenimiento de bienes públicos la elección racional consiste en no cooperar. Una consecuencia de ello, como explica Gordon Tullock (1971), es que el comportamiento no-cooperativo provoca que el cambio institucional a través de la acción colectiva, y particularmente las revoluciones políticas, sean difíciles de articular incluso en el caso de regímenes totalitarios y graves crisis político-económicas. El cálculo racional individualista propicia, de este modo, conductas que favorecen la estabilidad institucional en situaciones no Pareto-óptimas. Los mismos argumentos pueden aplicarse a la producción y difusión sociales de cualquier tipo de innovación, ya sea informal, o científica y tecnológica. La elección racional invita a decidir conservadoramente aun en situaciones tecnológicas claramente superables en sentido paretiano. Esto es un contrasentido desde la perspectiva teórica del equilibrio general.

La cooperación en el seno de los grupos, especialmente dentro de grandes colectivos de población, no está garantizada según el modelo estándar de racionalidad. Efectivamente, en la interacción de egoístas racionales en dilemas sociales y únicas jugadas no se producirá cooperación, aunque cooperar en la acción colectiva pueda resultar beneficioso para todos los jugadores como grupo. El éxito de la acción colectiva es un bien público (salvo que se pueda excluir de él a los que no se comprometen), y debido a su naturaleza, los agentes racionales con mentalidad cortoplacista tienen incentivos al remoloneo en el esfuerzo individual o la financiación de medidas a tomar. Si resulta inverosímil detectar y excluir a los *free-riders*, algo que ocurre, *cæteris paribus*, en marcos de información privada, o cuando los grupos son de gran tamaño y/o heterogéneos, el fracaso estaría asegurado, salvo que admitamos motivaciones prosociales, al estilo de un sostenido “entusiasmo militante” como el que plantea Konrad Lorenz¹⁶⁹. Pero este entusiasmo militante, de apariencia irracional, nos saca del paradigma del egoísmo psicológico y la racionalidad individualista propios de la teoría

¹⁶⁹ Para Lorenz (1978[1963]: 302-304) el “entusiasmo militante” es un instinto autónomo y diferenciado dentro del etograma de algunas especies de animales sociales, como la humana, que motiva para la cohesión grupal.

microeconómica y de la teoría de juegos ortodoxas. El vínculo del individuo con sus afines salta las barreras de los modelos de egoísmo racional, y se muestra como mecanismo eficaz para consolidar compromisos de cooperación que resultan favorables tanto para el individuo como para su grupo¹⁷⁰.

3.2.1.- Cooperación inestable en juegos de una única jugada

Los modelos de estática comparativa, enfocados al estudio de la interacción del momento presente (esto es, con tasas de descuento hiperbólicas y sin descuento del futuro), equivalen a juegos de una única jugada. En estos modelos la estrategia dominante en la acción colectiva de individuos egoístas es la de no cooperar, y por ello caen en las trampas del dilema del prisionero y otras tragedias colectivas. Los modelos de microeconomía estándar y teoría de juegos parten de estas premisas conductuales, y concluyen en la necesidad de arreglos contractuales vinculantes que acrecienten la credibilidad de los compromisos asumidos, así como, y especialmente, en el refuerzo de los mismos a través de jerarquías ejecutivas con poder para encauzar con incentivos el comportamiento oportunista (Alchian y Demsetz 1994[1972]; Williamson 1989). El grupo organizado se concibe entonces como ente político dotado de estructuras de poder especializadas en la supervisión, administración de incentivos y disuasión de los conflictos de interés inherentes a la acción colectiva. Esto resulta tanto más relevante cuanto mayor sea el tamaño poblacional, más complejas sus interrelaciones y menos transparentes desde un punto de vista informativo (Ostrom 1998). Como resaltan Gintis y Khurana (2008: 313), el diseño racional de un sistema de gobierno de la acción colectiva transforma supuestamente el problema de la no-cooperación en los dilemas sociales en «un problema político resuelto». En efecto, el tratamiento de la acción colectiva en los modelos de microeconomía estándar consiste en acuerdos reforzados por una tercera parte pero, tal y como sostienen Gintis y Khurana, las transacciones económicas e interacciones sociales más frecuentes en la vida cotidiana «se caracterizan por la *ausencia del refuerzo de esa tercera parte*» (*ibíd.*; cursivas en el original). Por esta razón muchos de los juegos de la vida real vienen a representar relaciones en las que desempeñan un papel crucial la confianza (*trust*) y la fiabilidad (*trustworthiness*) que aporten los jugadores, para lo cual se requiere disponer de información acerca del

¹⁷⁰ Para una explicación evolutiva de esta tesis véase Kanazawa (2004a).

otro jugador, sobre su reputación, capacidad y disposición para evitar ser explotado (Berg, *et. al.* 1995; Dasgupta 2000). Los agentes, de este modo, cooperan porque se necesitan mutuamente y además se da un cierto equilibrio disuasorio de posiciones de dominio.

En esta línea de análisis Robert J. Aumann (2005) ofrece una explicación estratégica de la cooperación en juegos de una sola jugada. El autor explica que el resultado de un juego se considera cooperativo si los jugadores eligen un resultado individual que, sin ser el mejor de todos los posibles para cada individuo en cuestión, sin embargo es la mejor alternativa que pueden *garantizarse* (Aumann 2005: 353). Cualquier otra opción verosímil conduciría a un equilibrio de Nash en el que todos los jugadores están peor. El principal problema decisivo en ambientes de información imperfecta es la existencia de equilibrios múltiples (Binmore 1994; Colman 2003). Por eso, en juegos de tirada única la cooperación requiere de un acuerdo previo entre los jugadores. Este acuerdo puede ser incumplido, dado que cooperar no es una estrategia óptima desde la perspectiva individual aunque sea el mejor resultado conjunto. Así, en el dilema del prisionero el resultado cooperativo surgirá si los jugadores se ponen de acuerdo en no confesar y ambos cumplen su palabra. Sin embargo, en este juego siempre existe la tentación de traicionar lo pactado, dado que la mentalidad individualista conduce a la paradoja racional del peor resultado de los posibles para ambos intervinientes (Schmid 2005b: 176-177). Cooperar, en este sentido, llega a ser teóricamente inverosímil a no ser que los individuos cuenten con algún incentivo de orden superior que les motive. Un ejemplo de ello puede ser la existencia de un tercero que refuerce el acuerdo y obligue a cumplir los compromisos asumidos a través de castigo al infractor (véase figura 3.1.)

Figura 3.1.- Matriz de pagos del juego Dilema del Prisionero

		Prisionero 1		
		Cooperar	Traicionar	
Prisionero 2	Cooperar	σ σ	γ δ	
	Traicionar	δ γ	β β	

La estructura genérica de los pagos del dilema del prisionero consiste en que $\delta > \sigma > \beta > \gamma$. El resultado cooperativo es (σ, σ) , que es mejor para ambos que (β, β) . No obstante, si ambos piensan que no existen garantías de que el otro jugador coopere, el resultado que obtendría el que sí cumple el pacto sería el peor de los posibles (γ) , mientras que su oponente obtendría el mejor de todos (δ) . Es por esto que si no existen refuerzos institucionales la estrategia elegida por los dos será traicionar (existe un único equilibrio de Nash, que es (β, β)), lo cual supone un resultado peor para ambos que el derivado de cooperar.

Obsérvese que Aumann define la cooperación como *lo mejor para todos* en contraposición a *lo mejor para uno mismo*, siendo este último, como se infiere, un resultado subóptimo desde la perspectiva *grupal*. Esto implica que si se trata de una interacción desarrollada por individuos racionales capaces de anticipar acontecimientos (o sea, realizada de forma voluntaria, consciente e intencional), entonces tendrán en cuenta en la elección de su estrategia que el mejor resultado no es *ganarlo todo* (primer óptimo individual), sino *ganar lo máximo compatible con el consentimiento de los demás* (segundo óptimo individual). De no ser así conseguirán, sin ser esta su intención, un tercer óptimo desde la perspectiva individual y grupal (un equilibrio no cooperativo, que no es óptimo paretiano individual ni grupal). Esta capacidad cognitiva de anticipación estratégica es lo que permite tomar en serio la oportunidad de un acuerdo. Aunque otros análisis de la cooperación, como el de Axelrod (1981; 1996) y Colman (2005), no requieren los supuestos de racionalidad ni inteligencia autoconscientes en los jugadores (como sería el caso de las bacterias, o incluso de programas de inteligencia artificial), sin embargo en el análisis de Aumann se está describiendo comportamiento humano, y en este sentido consciente de las coordenadas espacio - tiempo, y cognitivamente caracterizado por la disposición de algún tipo de capacidad para la inferencia de los pensamientos e intenciones ajenas. Obviamente, ese “descuento” de intenciones que lleva al equilibrio de Nash en el dilema del prisionero está basado en la independencia estratégica de los jugadores y en una creencia a destacar: el supuesto teórico del egoísmo motivacional. Ello implica que los individuos pretenden maximizar

su resultado individual sin tener en cuenta nada más que los resultados propios, y suponiendo que todos los demás jugadores actúan de ese modo. Esto supone admitir que los jugadores no están influidos, aunque sí condicionados, por las metas y los resultados esperados de los demás intervinientes (Myerson 1999). Tales presupuestos son característicos de situaciones competitivas, es decir, no-cooperativas. Pero aún asumiendo este patrón de racionalidad, los jugadores humanos pueden, dado que disponen de una teoría universal de la mente, haber aprendido de su experiencia pasada y así, anticipando la repetición de fracasos, modificar su estrategia de defección y cambiarla por una estrategia inteligente de cooperación. Un resultado similar puede ser fruto de procesos de evolución a largo plazo en organismos no capacitados para el aprendizaje y la memoria (Colman 2005; Alchian 1950). En la figura 3.2. se muestra esta posibilidad de la anticipación estratégica con resultados favorables para el conjunto y sin que se requiera identificación intersubjetiva, vínculo simpatético u otras expresiones de una mentalidad práctica de equipo.

Figura 3.2.- Matriz de pagos del juego G

		Jugador B		
		Consentir	Castigar	
Reparto equitativo	50	50	0	Jugador A
	Reparto codicioso	99	1	

Fuente: elaboración propia basada en Aumann (2005: 353)

Nota: el jugador *B* aceptaría el reparto (99, 1) dentro de los cánones convencionales de racionalidad, si bien él preferiría el reparto equitativo (50, 50). Un planteamiento conductual de tipo egoísta conducirá al equilibrio de Nash (99, 1). No obstante, una cantidad destacada de los experimentos llevados a cabo en laboratorio no concluyen de este modo, debido a que el jugador *B* generalmente manifiesta preferir perder esa ganancia de 1 en lugar de aceptar un reparto tan desequilibrado (Güth, *et. al.* 1982; Hoffman, *et. al.* 1998; Oosterbeek, *et. al.* 2004). Cuando existen jugadores que aplican venganza ante propuestas de reparto “injustas” la interacción conduce a un resultado aparentemente irracional (0, 0), que es coherente con una estrategia emocional de compromiso (Frank 1988), acorde con una interpretación microeconómica evolutiva como la de Frank (1994: 244-254).

En el ejemplo propuesto (el juego *G*) existen dos jugadores *A* y *B*. El primero es quien inicia el juego proponiendo un reparto, que puede ser equitativo (50,50) o codicioso a favor de sí mismo (99,1). El segundo jugador tiene la capacidad de dar por bueno el reparto propiciado por *A*, u oponerse a él aplicando un castigo por el cual el resultado de la distribución sería (0,0), el peor para ambos. La estrategia de egoísmo

maximizador por parte de A sería ofrecer (99,1), pero como descuenta que B no aceptará dicho reparto (aplicará la respuesta de castigo aunque pierda 1 unidad de ganancia), por ello la estrategia racional en esta matriz de incentivos será cooperar, repartiendo equitativamente las ganancias¹⁷¹. Aumann, no obstante, subraya que este resultado aparentemente cooperativo no es un resultado de equilibrio, ya que cooperar de este modo no proporciona la satisfacción de los mejores intereses de A . Esto es así porque su intención predominante será (como en el juego del dilema del prisionero) la de apropiarse de la mayor cantidad posible de ganancia. En el juego G que hemos explicado existe posibilidad de una disuasión por parte de B para evitar un reparto no equitativo de las ganancias. Aún así, no se produce ningún equilibrio de Nash. En otros juegos, como el dilema del prisionero, la disuasión (en forma de represalia descontable) ante las conductas no cooperativas no es posible. De este modo, el resultado cooperativo no puede considerarse un resultado de equilibrio, sino el resultado del acuerdo, ya sea éste expreso o tácito¹⁷².

En interacciones sin más horizonte temporal que el momento presente los individuos egoístas cooperan por contrato, y esto sólo es factible en contextos en que los acuerdos son ejecutables, es decir, cuyo cumplimiento está garantizado por un tercero (Estado, jerarquía organizacional, árbitro imparcial, etc.) con poder para hacer cumplir

¹⁷¹ Téngase en cuenta que se está admitiendo de forma implícita que B considerará injusto e indignante el reparto (99,1), y que por lo tanto aplicará el castigo aunque le cueste perder la unidad que se le ofrece con tal de que A pierda también toda ganancia posible. Se está pues asumiendo los rasgos conductuales propios de los resultados de experimentos del tipo juego del ultimátum, en los que los jugadores cooperan aun en situaciones de dilema del prisionero, y que sólo dejan de cooperar, en juegos finitos, en la última o últimas jugadas (Ones y Putterman 2007; Plott 1996; Ledyard 1995; Sally 1995). Existe, por otra parte, una abundante literatura teórica y experimental que apoya la tesis de que la posibilidad de aplicar un castigo al no-cooperador permite la evolución de la cooperación grupal tanto a muy largo plazo como en contextos finitos. De este modo, la presencia conductual de mecanismos desencadenadores de la venganza (*strong reciprocation*), ya sean activados por juicios relativos a la justicia, coherencia con determinados preceptos morales, cumplimiento de las normas sociales o por emociones innatas, sería un factor efectivo de refuerzo de la cooperación. Véanse, por ejemplo, estudios sobre el sentimiento de justicia y las expresiones de indignación, y su presencia en el comportamiento en juegos del ultimátum llevados a cabo por primates no humanos en Brosnan y De Waal (2003), y en De Waal y Luttrell (1988). En esta misma línea, trabajos relacionados con la conducta de castigo al oportunista en individuos humanos pueden consultarse en Ones y Putterman (2007); Masclet *et. al.* (2003); Fehr y Gächter (2000a, 2000b y 2002); y Sefton, *et. al.* (2002). Dos trabajos experimentales adicionales que demuestran la existencia de preferencias favorables a una justa distribución de los pagos en humanos (aversión a la inequidad) son los de Bolton y Ockensfels (2000), y Fehr y Schmidt (1999).

¹⁷² Coincide este planteamiento, en líneas generales, con las tesis defendidas por Gauthier (1994[1986]) en *La moral por acuerdo*.

coercitivamente los compromisos asumidos, establecer normas y obligar a cumplirlas¹⁷³. No obstante, es preciso tener en cuenta que pueden existir contratos incompletos por su propia naturaleza (como la relación principal-agente), y esta tipología complica aún más el análisis, dado que la existencia de equilibrios múltiples lleva a una incierta (y a veces inoperante) definición previa de compromisos explícitos, difíciles de establecer con precisión, lo cual hace de estos contratos una fuente de costosa controversia aun admitiendo la solución hobbesiana¹⁷⁴.

Así pues, en modelos de interacción estáticos y “sin futuro”, o con una tasa de descuento del futuro elevada, la cooperación no es un equilibrio estable salvo que los contratos estén reforzados por una autoridad central e imparcial. En este marco relacional y de incentivos es donde aparecen la necesidad de tener en cuenta la historia y la reputación, la memoria emocional, los arreglos institucionales para reforzar la credibilidad de los compromisos, y las expectativas de un futuro sobre el que decidir racionalmente. Pasemos a explicarlo.

3.2.2.- Cuando se cree que la interacción es inagotable: cooperación en superjuegos

¿Por qué se producen estrategias de cooperación estable en el tiempo dentro de estos contextos de velado enfrentamiento? ¿Por qué el egoísmo racional no conduce a un resultado de suma negativa a corto plazo? Douglass C. North ha resaltado que las conclusiones más desalentadoras de Olson, díganse la trampa del dilema del prisionero y la tragedia comunal, son los resultados lógicos de encajar el supuesto del individualismo maximizador en un marco de análisis estático, en el que el juego sólo tiene una jugada (North 1993: 25)¹⁷⁵. No obstante, argumenta que una buena parte de las interacciones no son de ese tipo, por lo que se requiere otro escenario analítico más realista. Este escenario alternativo nos saca de la estática y nos sitúa en el análisis dinámico, a incluir la variable tiempo en el modelo, tener en cuenta tasas de descuento, asumir expectativas acerca de la motivación de los jugadores y de su influencia en los

¹⁷³ Aumann (2005: 354) cita a propósito una frase del Talmud que dice: “Ruega por el bien del gobierno, pues sin su autoridad el hombre se tragaría vivos a los demás hombres”, una sentencia no muy alejada del aforismo hobbesiano de los lobos.

¹⁷⁴ Véase Boulding (1971b), y Milgrom y Roberts (1993: cap. 5).

¹⁷⁵ Este es el marco conductual en el que se desarrollan las transacciones impersonales de un mercado con individuos anónimos, y por tanto desconocidos entre sí (Gintis y Khurana 2008: 307).

pagos futuros, y expandir el juego para transformarlo en un superjuego (G^∞). Dentro de este contexto se entiende el razonamiento expuesto por Mises acerca de la cooperación empírica cuando no existen vínculos afectivos, información acerca de la fiabilidad, benevolencia o altruismo:

La cooperación social y su corolario, la división del trabajo, sirve mejor a los intereses egoístas de los individuos afectados que el aislamiento y el conflicto. Cada paso hacia una cooperación pacífica ocasiona una ventaja inmediata y discernible para todos. Los hombres cooperan y anhelan intensificar la cooperación porque desean perseguir sus intereses egoístas. Los sacrificios que el individuo hace para el mantenimiento de la cooperación social son sólo temporales. Si se abstiene de las acciones antisociales que podrían darle pequeñas ganancias inmediatas, gana mucho más por las ventajas que recibe de la cooperación pacífica (...) En el largo plazo, no hay conflicto irreconciliable entre los intereses egoístas correctamente entendidos de los individuos y los de la sociedad. (Mises 2005[1944]: 217-218).

Desde el campo de la sociobiología hallamos declaraciones semejantes a la enunciada por Mises. Por ejemplo, en un texto de Edward O. Wilson éste sostuvo que:

Los seres humanos parecen ser lo suficientemente egoístas y calculadores como para ser capaces de una armonía y equilibrio sociales muy grandes. Esta afirmación no es autocontradictoria. El verdadero egoísmo, si es obediente a todos los principios de la biología de los mamíferos, es la clave de un contrato social casi perfecto. (Wilson 1983: 156-157).

Mises y Wilson, cada uno desde su campo, están esbozando una teoría de la cooperación entre individuos egoístas e inteligentes, sin necesidad de identificación intersubjetiva ni de vinculación afectiva, en el sentido de que son capaces de comprender que conseguir toda la ganancia extra posible a corto plazo es un “mal negocio” a largo plazo.

La visión predominante en la microeconomía es la de una cooperación que, en ausencia de información perfecta, para existir ha de estar reforzada por instituciones racionalmente diseñadas. Esta interpretación hobbesiana del contrato social plantea, no obstante, algunas incertidumbres. Así, esa tercera parte implicada en el juego (el Leviatán hobbesiano) también es un jugador como los demás, potencialmente calculador y oportunista y, por lo tanto, capaz de traicionar su misión de agencia guiado por intereses particulares. Reconocer esto supone reconocer que se necesitan nuevos diseños de incentivos y controles en el entramado institucional, un cierto equilibrio de

poderes *à la Montesquieu* que evite la arbitrariedad y las alianzas en contra del propio sistema cooperativo. Pero no hay certeza de que esto ocurra: «de momento nadie sabe cómo crear tal entidad» (North 1993: 82). Se trata del viejo problema de controlar al controlador. Douglass North lo ha especificado del siguiente modo:

Aquí precisamente es donde se encuentra el dilema fundamental del desarrollo económico. Si no podemos lograrlo sin el Estado, tampoco podemos obtenerlo con él. ¿Cómo esperar que el Estado se conduzca como una tercera parte imparcial? (...) En pocas palabras, si el Estado posee la fuerza coercitiva, entonces quienes rijan el Estado usarán esa fuerza en su propio interés a expensas del resto de la sociedad. (North 1993: 81-83).

La alternativa consiste en que cuando el juego es un superjuego no es necesario un tercer agente cuya función consista en reforzar la cooperación con el monopolio exclusivo del poder. Es suficiente con asumir que existirán represalias ciertas en interacciones futuras por parte de quien ha empezado cooperando y ha recibido defección como respuesta. No se requiere un Estado instrumental. Esto nos introduce en la conducta de reciprocidad en juegos de más de una jugada. ¿Es realmente suficiente el ejercicio de la reciprocidad para que la cooperación pueda evolucionar frente a la defección en juegos del tipo del dilema del prisionero y los dilemas sociales al uso? Reflexionemos algo más acerca de este tema.

3.3.- *Homo reciprocans*: la evolución de la cooperación mediante la estrategia del “toma y daca”¹⁷⁶

En la ciencia económica se identifica la reciprocidad con el intercambio motivado por la maximización de utilidad por parte de dos agentes. Sin embargo, cuando entre los individuos se da simplemente el «intercambio mercenario de buenos oficios» que menciona Adam Smith (1997[1759]: II, ii, 3. p. 186), surge lo que Avner Greif (2004: 10) ha llamado “problema fundamental del intercambio”, es decir, la paradoja del egoísmo. Dicha paradoja se puede formular mediante las siguientes interrogantes: ¿Qué incentivo puede llevar a los egoístas a cooperar, cuando la estrategia racional es no hacerlo? ¿Qué puede impulsar a individuos egoístas a actuar en pos de un beneficio conjunto y no para su propio beneficio individual e inmediato? La mayor parte de los modelos más conocidos al respecto versan sobre la conducta de

¹⁷⁶ Toma y daca es la expresión española de los circunloquios en inglés “tit for tat”, que utiliza Axelrod (1996), o su equivalente “an eye for an eye, and a meal for a meal” de Sigmund y Nowak (1997).

reciprocidad, entendida genéricamente como la relación caracterizada por una tendencia conductual a «entregar bienes en respuesta a los bienes recibidos, y también a cobrar venganza (por los males)» (Kolm 2000: 1; paréntesis añadido). Se trataría pues de una estrategia adaptativa de cooperación condicional, por la que existe un compromiso pre-racional a largo plazo con una regla distributiva de equidad¹⁷⁷.

Adam Smith consideró, en los fundamentos de su modelo de sociedad comercial, que el *Homo sapiens* es un animal cooperativo porque instintivamente tiende a dividir el trabajo y a intercambiar bienes. Asimismo, Smith refirió la posibilidad evolutiva del propio interés a la existencia de una psicología capacitada para la justicia (respetar lo ajeno aunque sea en contra del propio interés inmediato), y un marco institucional que favorezca el incentivo a comportamientos justos. Utilizando una terminología moderna, diríamos que Smith está apuntando directamente a la reciprocidad como base epigenética de lo que él llama “sociedad civilizada”, basada en el intercambio justo (y no en tomar las cosas por la fuerza) de todo tipo de gestos, acciones y objetos dentro de un orden social estable¹⁷⁸.

Obviamente, el intercambio comercial es una parte de ese orden social cooperativo (en este caso una reciprocidad explícita, actual, contractual y previamente convenida), y la economía estudia todo lo que se deriva de la existencia de ese instinto que impele al intercambio recíproco en mercados más o menos formales. Cosmides y

¹⁷⁷ Intercambiar y comerciar está presente en el género homo al menos desde hace 1,2 millones de años (*Homo erectus* o *ergaster*) y se trata de una adaptación instintiva. El mercadeo no se aprende, se sabe; lo cual va asociado a un sentido innato de la posesión y de la propiedad (Rossano, *et. al.* 2011). El mercadeo es intrínseco a hacer herramientas y tener lenguaje, y junto con el aprendizaje forman la tétada de la cultura. Existen evidencias arqueoantropológicas que demuestran que el yacimiento de Olorgesailie (Kenia) es el primer taller conocido de producción masiva de útiles líticos. Los materiales para su fabricación proceden de 14 canteras, algunas situadas a 18 Km., lo que indica la existencia de división del trabajo, así como interacción intra e inter tribal de diferentes poblaciones de *ergaster*, tanto para el acceso a las diferentes materias primas como para adquirir los útiles del taller. Desde el comienzo del *Homo sapiens* estos dos elementos (talleres de producción masiva, como el de Blombos Cave en Cape Town, Sudáfrica; y presencia de útiles en regiones distantes) se hallan presentes, si bien en mayor escala y a mayores distancias. Esto permite inferir, tal y como han resaltado Henshilwood, *et. al.* (2011: 220), que «en la evolución de la relativamente compleja psicología humana, ya se dio desde el principio la capacidad conceptual para acaparar provisiones, combinarlas y almacenar mercancías que eran utilizadas para mejorar las prácticas sociales».

¹⁷⁸ Véanse al respecto el análisis de Smith (1997[1759]: II, ii, 3), así como la reflexión de Ridley (1996: cap. 2) acerca de la naturaleza comercial de la especie humana. La base epigenética queda demostrada en la nota anterior, ya que el intercambio fue una actividad presente en el género *Homo* antes del *Homo sapiens*.

Tooby defienden que la reciprocidad es un producto de la selección natural, proceso que ha generado mecanismos fisiológicos y psicológicos que motivan al intercambio porque tal intercambio recíproco, en lugar de la agresión, la extorsión y el saqueo, favorece a los individuos que lo practican en el plano biológico (Cosmides y Tooby 1992: 169)¹⁷⁹. Esta hipótesis concilia bien con el hecho constatado de que la reciprocidad es un patrón de comportamiento universal en la especie humana [Henrich *et. al.* (2005 y 2001); Pinker (2004: 642-664); Pierce y White (1999); Hoffman, *et. al.* (1998); Nicholson (1997); Wilson (1993: 60)], y puede apreciarse también en otras especies de animales sociales¹⁸⁰. Algunas formas de simbiosis *intra e inter* específicas pueden caer dentro de un concepto amplio de reciprocidad, si bien la práctica de la reciprocidad como tal exige ciertas capacidades cognitivas, igual que la cooperación. Así, algunos enfoques contemporáneos denominan cooperación a conductas que encuadraríamos en la simbiosis con reciprocidad elemental, dado que prescinden de los factores psicológicos y motivacionales, y explican la cooperación sólo como proceso (no-intencional) constructor de estructuras físicas y sociales de los organismos¹⁸¹.

Dentro del enfoque de los modelos de reciprocidad, el intercambio económico vendría a ser una modalidad específica instintiva del establecimiento de obligaciones recíprocas. En el caso humano este acto viene mediado por la intención consciente y la comunicación lingüística. Se trataría pues de una modalidad de comportamiento situada en la frontera del altruismo débil y el egoísmo impuro.

¹⁷⁹ La norma de reciprocidad a la que se refieren los sociólogos y antropólogos es una manifestación de la ultrasocialidad biológica de la especie humana (Richerson y Boyd 1998), y pudo ser seleccionada en el etograma a través de un proceso coevolutivo que benefició al linaje de los individuos que la practicaron (Hoffman, *et. al.* 1998: 380). Naturalmente, en este proceso pudo jugar un papel decisivo la regulación político-moral de las relaciones fuera del núcleo familiar. Ha de tenerse en cuenta que la cultura, fenómeno consustancial a la especie humana, cuenta con una trayectoria histórica de entre 120.000 y 150.000 años. La expresión social - institucional de la norma filogenética de reciprocidad evita el mundo hobbesiano y la lucha de gladiadores por los recursos escasos y, de este modo, los individuos y los grupos que la practicaron fueron capaces de evitar la trampa de los bienes comunales y los juegos de suma cero (Sigmund y Nowak 1997).

¹⁸⁰ Por ejemplo, los comportamientos de ayuda mutua estudiados en vampiros tropicales y delfines (Ridley 1996: 62-63; 160-164), el altruismo y la reciprocidad observados en primates, delfines cetáceos y paquidermos (De Waal 1997: 57, 75 y 85), o los frecuentes casos de solidaridad recíproca catalogados en aves (Weismann 1994).

¹⁸¹ Por ejemplo, los genes, las células o los órganos vitales de un organismo actúan *como si* cooperasen (Mayr 2000); sin embargo no se puede afirmar que sean unidades de análisis dotadas de comportamiento. Así, las bacterias pueden establecer relaciones simbióticas con ciertos organismos, pero nos parece una exageración proponer que actúen mediante patrones racionales o intencionales de reciprocidad, y mucho menos de cooperación. La reciprocidad y la cooperación son conductas propias de organismos dotados de una psicología evolucionada.

El sociólogo David Gouldner presenta una definición de la reciprocidad mínima según la cual ésta consiste en un patrón de comportamiento (o estrategia de acción) basado en dos reglas simples aplicadas por los individuos (Gouldner 1960: 171): (a) ayudar a aquellos que les han ayudado; y (b) no rechazarles o atacarles (o sea, no practicar defección) mientras que dicha ayuda siga produciéndose. El autor presenta la reciprocidad humana como una norma universal que tiene expresiones diferentes dependiendo de aspectos de índole cultural, pero que, en cualquier caso, resulta esencial para la estabilidad de los grupos y los sistemas sociales¹⁸².

En biología este comportamiento de entregas y ayudas recíprocas se ha denominado tradicionalmente simbiosis, una de cuyas manifestaciones es el mutualismo, o simbiosis mutuamente beneficiosa (Milner 1995: 230). Uno de los ejemplos más citados en la literatura sobre simbiosis es el de la coevolución adaptativa entre las acacias tropicales de América Central y las hormigas de las acacias que explica Richard Milner (1995: 26). Ambas especies intercambian bienes y servicios: la colonia de hormigas vive en la corteza de la acacia y se alimenta de los nutrientes que la acacia produce. “A cambio” la colonia de hormigas actúa como un sistema inmune repeliendo y atacando a los insectos filófagos que habitan la región, que acabarían con la vida del árbol en cuestión de días ya que estas acacias no están protegidas por hojas o savia tóxicas que repelan a los parásitos.

Dentro del modelo biológico de reciprocidad natural, es posible practicar la cooperación en diversos niveles de intensidad y complejidad. Dichos estándares de cooperación fueron analizados inicialmente por Konrad Lorenz (1978[1963]: caps. 8-11), quien los categorizó en cuatro apartados:

- a) Lo que el autor denomina la *multitud anónima*, en la que los individuos viven muy próximos unos de otros y se mueven como grupo a través de comportamientos de tipo contagioso o mimético (*herd behavior*). Esta organización es típica de ciertas especies de peces, aves e insectos, algunas de

¹⁸² Planteamientos dentro del campo de la economía como Hoffman, *et. al.* (1998), o Fehr y Gächter (2000b), resaltan los elementos premio/castigo – beneficio/coste como respuestas estereotipadas a las conductas de ayuda/daño respectiva y previamente recibidas.

mamíferos y muchas de invertebrados, como los cefalópodos, que se asocian en grandes grupos gregarios en los que la cooperación -una cooperación marcadamente pasiva- consiste básicamente en la presencia física dentro del grupo, que se convierte en una defensa contra los depredadores y un modo más eficaz de obtener alimento.

- b) La *sociedad sin amor*, que consiste en sociedades, a veces bastante duraderas, en las que los individuos cooperan muy básicamente (p.ej. en la defensa de un territorio o en la construcción de anidamientos), aunque de una forma más activa que la multitud anónima (no se trata solo de la mera presencia gregaria), y sin existir lazos interpersonales ni con el grupo ni con individuos específicos. Se trata de una simbiosis en un sentido estricto. Es el caso de las formas de vida, por ejemplo, del lagarto verde del sur de Europa y las cigüeñas europeas.
- c) Sociedades endogámicas, como las *superfamilias* que agrupan a varias generaciones de ratas, lobos, insectos sociales, etc., en las que se da un reconocimiento individualizado pero una cohesión de tipo clánico: el individuo es uno más en el clan, que viene delimitado por *marcadores colectivos* como son las feromonas u otros rasgos diferenciadores del linaje. En estas sociedades endogámicas la cohesión inclusiva se refuerza a través de comportamientos beligerantes hacia los *outsiders*. Aquí sí se producen comportamientos de cooperación y altruismo, si bien se trata de un altruismo hacia un marcador, sea quien sea el individuo que lo porta.
- d) Y lo que Lorenz llama *grupo* propiamente dicho, caracterizado por ser un agregado de vínculos individualizados (no colectivizados, como en el caso anterior de las superfamilias), propio de algunas especies de aves y de primates cognitivamente avanzados, como los humanos:

La formación de un grupo verdadero presupone que los individuos son capaces de reaccionar selectivamente a la individualidad de sus vecinos o compañeros (...) es condición *sine qua non* para la formación de un grupo la *identificación personal* del compañero (...) identificación que se realiza, claro está, individualmente. (Lorenz 1978[1963]: 188; cursivas en el original).

¿Por qué ciertas especies han evolucionado en sus pautas de comportamiento de tal modo que sus individuos se agrupan y cooperan mientras que en otras no ocurre esto? La explicación que ofrece Dawkins a este fenómeno biológico es un argumento genuinamente económico:

... si los individuos viven juntos en grupos, sus genes deben obtener de la asociación un beneficio mayor de lo que invierten en ella (...) En general, las asociaciones en beneficio mutuo evolucionarán si cada socio obtiene más de lo que aporta. (Dawkins 1993[1976]: 217 y 328).

Según la visión reduccionista de Dawkins, lo que Lorenz llama *vinculación* personal a una identidad individual o grupal es un asunto secundario que esconde una contabilidad inconsciente de coste-beneficio. Dawkins huye de valoraciones psicológicas en la acción de los organismos y se centra exclusivamente en la utilidad instrumental que aporta la relación recíproca a cada uno de los individuos: «Aquí no hay altruismo, sólo existe la explotación egoísta por parte de cada individuo a costa de los demás» (Dawkins, 1993[1976]: 220). Siendo esto así, la teoría de los grupos que plantea Dawkins es de carácter individualista, y excluye de forma expresa el altruismo, entendido éste como contribución a aumentar el *bienestar* de otro animal (identidad genética diferente) a expensas del propio¹⁸³: Las “máquinas de supervivencia” de Dawkins en realidad nunca son altruistas, y cuando parece que lo son es porque están maximizando la presencia de sus genes en otros organismos además del “suyo”. Expliquemos el contenido de esta cooperación por reciprocidad entre egoístas.

3.3.1.- Reciprocidad directa: respuestas de gratitud y represalia

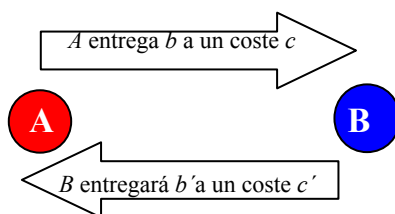
El intercambio económico es una modalidad de simbiosis por reciprocidad directa, que es previamente pactada y más o menos inmediata (Hirshleifer 1978: 241), en la que el agente persigue el objetivo de intercambiar apoyos supeditados a que el valor subjetivo de lo recibido supere al del coste de lo entregado (Nowak, *et. al.* 1995: 77). En la reciprocidad directa el individuo *A* soporta un coste *c* para que otro individuo *B* obtenga un beneficio *b*, y simultáneamente *B* soporta un coste *c'* para que el individuo *A* obtenga un beneficio *b'*. Ambos son “altruistas recíprocos” puesto que los dos

¹⁸³ El bienestar según Dawkins (1993[1976]: 5) es una metáfora de las oportunidades de supervivencia, por muy insignificantes que sean.

soportan un coste; y además, en el caso de organismos cognitivos capacitados para la intersubjetividad ambos son altruistas psicológicos, ya que de algún modo su intención es satisfacer los deseos de otro individuo. Al ser cooperadores recíprocos cada entrega viene condicionada por la regla “si yo coopero ahora, tú cooperarás después” (ese “después” suele ser preestablecido, y con frecuencia se trata del mismo momento temporal). Esta es en esencia la teoría de la cooperación de Trivers (1978[1971]), con la particularidad de que los bienes entregados se pactan, e igualmente el tiempo que transcurre entre el dar y el recibir (véase figura 3.3.) En sus fundamentos la teoría económica parte de que los individuos humanos propenden a cooperar condicionadamente, y que lo hacen motivados por la conveniencia personal. Así, utilizando palabras de Wicksteed:

... si hago una cosa no porque tenga un interés directo en ella, sino porque lo desea alguna otra persona, quien a su vez hará lo que yo quiera o me proporcionará los medios de obtenerlo, entonces estoy cumpliendo sus deseos *como medio de satisfacer los míos*. Al cumplir los suyos, estoy indirectamente satisfaciendo los míos. En esto consiste la naturaleza de las relaciones económicas, y el objeto propio de la investigación económica es el mecanismo o articulación de todo el complejo de estas relaciones. (Wicksteed 1963[1914]: 3; cursivas añadidas).

Figura 3.3.- Esquema de la reciprocidad directa



La lógica de la reciprocidad directa, desde una perspectiva económica, está respaldada por el principio de la utilidad marginal decreciente. De este modo, en el momento del intercambio $U_A(b') > U_A(c)$; y $U_B(b) > U_B(c')$. Esto es así porque c es un “excedente” que aporta baja utilidad al individuo A en el presente, mientras que b' puede resultarle imprescindible para sobrevivir, o altamente útil, en el momento que lo reciba. El mismo razonamiento se aplicaría al individuo B . Si el *do ut des* no es cuasi-instantáneo, sino que media tiempo $t > 0$ entre la entrega y la reciprocidad, entonces intervendrá una tasa de descuento, de modo que el individuo racional actuaría como si $[U_A(b') / (1 + r)^t] > U_A(c)$, donde r es la tasa de descuento, y t el intervalo de tiempo que transcurre entre dar y recibir. En este caso, b' y c pueden ser el mismo bien en dos momentos temporales distintos.

Fuente: elaboración propia.

En la teoría de juegos estándar se parte del axioma motivacional y epistemológico *no-tuista*, un presupuesto, por otra parte, clásico en la historia de la teoría económica del intercambio de mercado¹⁸⁴. En este marco analítico no-tuista Axelrod ha analizado la estrategia más robusta en torneos de algoritmos computarizados enfrentados en interacciones repetidas de un superjuego¹⁸⁵. La estrategia victoriosa de los torneos fue la ideada por el psicólogo y teórico de juegos Anatol Rappoport, y consiste en cooperar siempre en la primera ronda del juego, y en sucesivas jugadas posteriores hacer lo que la contraparte hizo en la ronda anterior (Axelrod 1996: 30). Este patrón de comportamiento fue bautizado con el nombre “Tit for tat” (*tft*), que puede traducirse por “toma y daca”, “donde las dan las toman”, o “pagar con la misma moneda”. Axelrod analiza cómo esta estrategia cooperativa pudo “evolucionar” en un mundo de organismos independientes y egoístas generando un poderoso sistema espontáneo de autocumplimiento. Téngase en cuenta que el autor descarta el altruismo original por ser una estrategia abocada al fracaso evolutivo, aunque curiosamente *tft* consiste en comenzar el juego “dando”, lo que podría entenderse como una forma de altruismo *à la* Trivers. No obstante, el altruismo indiscriminado fracasa sistemáticamente como estrategia en los torneos cuando se enfrenta a estrategias egoístas de forma reincidente¹⁸⁶.

El núcleo de la hipótesis consiste en que las interacciones bilaterales evolucionarán cooperativamente siempre que la tasa de descuento del futuro sea lo suficientemente baja, y que se dé además una expectativa altamente probable de que los jugadores vuelvan a encontrarse en futuras interacciones¹⁸⁷. Aunque Axelrod explica que para que esto ocurra no se requiere conciencia ni racionalidad deliberativa en los jugadores (es decir, no se requiere elección propiamente dicha, basta con pautas de

¹⁸⁴ Como explica Kolm (2000a: 8 y 17) este *no-tuismo* figura literalmente de forma destacada en las obras de Wicksteed y Pareto. También en la axiomática inicial de los modelos teóricos de Jevons y Edgeworth, y en la teoría de juegos de Neumann y Morgenstern.

¹⁸⁵ En un primer torneo compitieron 14 estrategias, que más tarde fueron ampliadas a 62. En cada ronda una estrategia se enfrentaba a todas las demás, incluida una copia de sí misma. Esto conlleva un total de 225 (15 x 15) y 3.969 (63 x 63) partidas en cada torneo. Superadas las 200 rondas se sumaron puntuaciones y se estableció la clasificación final.

¹⁸⁶ Esta conclusión se acomoda bien a otras visiones de la interacción social. En general tanto la teoría de juegos, como la economía y la biología evolutiva consideran el altruismo una regla de comportamiento poco realista.

¹⁸⁷ Este es el llamado *Folk Theorem* de la cooperación en el modelo estándar de racionalidad (Fudenberg y Maskin 1986).

comportamiento deterministas), no obstante, el resultado cooperativo de las entidades programadas con *tft* sería equivalente a la elección racional de organismos cognitivos capaces de descontar correctamente el futuro en su función de utilidad. La lógica de la selección natural de los algoritmos en competencia permite entonces concebir el modelo de elección racional a modo de “como si” ésta existiera realmente en organismos que sí son conscientes y calculadores.

¿Cómo puede haber evolucionado este patrón de comportamiento social? La clave explicativa está en que el individuo se integre en grupos y redes de organismos que practican *tft*. En efecto, según los resultados de las simulaciones computarizadas, la estrategia *tft* puede ser una *estrategia evolutivamente estable* (EEE) si se arraiga en grupos de individuos *apiñados* en los que una masa crítica suficiente actúa según dicho patrón conductual¹⁸⁸. Esto quiere decir que *tft* es capaz de defenderse a sí misma con éxito frente a la invasión de otras estrategias no-cooperativas, o menos cooperativas. El que los agrupamientos de cooperadores condicionales sean limitados y relativamente *pequeños* es una exigencia de carácter instrumental, pues se requiere que los jugadores tengan una alta probabilidad de volver a interactuar en el futuro. Igualmente es una exigencia por razones de tipo cognitivo, ya que es imprescindible que los individuos recuerden y reconozcan bien a los cooperadores y defectores de anteriores jugadas para aplicarles la regla de reciprocidad (volver a cooperar -con los cooperadores-, o aplicar defección o eludir volver a jugar -con los defectores-). De no ser así, las estrategias oportunistas externas al apiñamiento, o la aparición de mutantes internos, invadirían el grupo y lo conducirían a la degeneración de su *ethos* relacional (y como consecuencia, dejarían de darse las condiciones para que pueda evolucionar la cooperación en el enclave cooperativo). Esta capacidad discriminatoria es necesaria para que la cooperación recíproca se asiente en el grupo y sobreviva, y ello es así porque si resulta improbable que los individuos vuelvan a repetir interacción, entonces el cooperador no podrá represaliar la defección previa de la que haya sido objeto. En condiciones de

¹⁸⁸ Axelrod (1996: 31) se refiere a los enclaves de cooperación aludiendo a «pequeños grupos apiñados que fundamentan su cooperación en el principio de reciprocidad». El enclave es un equilibrio social evolutivo producido por la interacción de estrategias individuales de comportamiento. Algunas de las características de los organismos más propicias para el surgimiento de enclaves son disponer de buena memoria (p.ej. para asociar rostros u otros rasgos identificatorios a comportamientos pasados), capacidad limitada de establecer relaciones estables más allá de cierto umbral cuantitativo, y un notable sedentarismo relacional.

pluralidad de estrategias simultáneas los individuos *tft* que carezcan de buena memoria, o que se desenvuelvan sin cautelas en ambientes relacionales despersonalizados o de interacciones únicas, son obvios candidatos a ser víctimas de la explotación de otros. Una consecuencia de esto es que cooperación no sería una EEE, y por ello tendería a desaparecer. Así, sólo en contextos de dilema del prisionero repetido puede evolucionar la cooperación basada en *tft*, y la condición para que esto ocurra es que la probabilidad p de que se vuelva a producir un encuentro entre los mismos dos individuos supere la relación genérica coste/beneficio expresada en la anterior figura 3.3., es decir, que:

$$p > c/b^{189}$$

Obsérvese que la descripción matemática de la desigualdad formulada por Nowak guarda relación con la regla de Hamilton [$b \cdot r > c \Rightarrow r > (c/b)$]. La parte derecha de la desigualdad es equivalente en ambos casos, y la parte izquierda representa, también en ambos casos, una probabilidad: la de volver a encontrarse en el futuro en el caso que estamos explicando, y la de que ambos organismos sean portadores del mismo ADN en la regla de Hamilton.

Lógicamente, si los emparejamientos son aleatorios p tiende a disminuir con el aumento del tamaño de la población (n), por lo que, en contextos de elevada incertidumbre debidos a una demografía muy dinámica o frecuentes mutaciones, la manera que tienen los cooperadores de protegerse frente a la defección es reduciendo su “tasa de generosidad”, o de “equitatividad”, (medida como el coste personal -p.ej. en *fitness*- de lo que entregan a desconocidos; es decir, el coste c de entregar b), o incluso evitando interactuar con aquéllos que no se conoce (lo cual equivale a que $c = 0$)¹⁹⁰.

¹⁸⁹ Para la demostración del teorema véase Nowak (2006a: 1.561). Entiéndase que el coste del que aquí se habla (c) hace referencia al coste soportado por el cooperador (medido en bienes o servicios, que en definitiva se pueden traducir en eficacia genética) por entregar algo a otro individuo, quien se beneficia de ello en b (medido en bienes y servicios, o sea, en eficacia genética). La relación c/b es una relación comparativa entre la eficacia genética que pierde el cooperador respecto a la eficacia genética que otro gana. En términos económicos este cociente consiste en una tasa de intercambio que representa el cociente de las utilidades marginales respectivas, derivadas de la transferencia unilateral de un bien sin que exista reciprocidad inmediata, sino una expectativa de reciprocidad futura.

¹⁹⁰ Zak, *et. al.* (2007) diferencian entre altruismo y generosidad (*generosity*). Altruismo consiste en ayudar a otro individuo soportando un coste por ello, mientras que generosidad -que aquí traducimos con el término más preciso de equitatividad, que utilizaremos en lo sucesivo- es la liberalidad en lo que se entrega, u ofrecer al otro más de lo que este espera o necesita. La generosidad alude pues a la gradación del altruismo: «Por ejemplo, uno puede entregar a una persona sin hogar 25 céntimos

Siendo $p = 1/(n-1)$, cuando $n \rightarrow \infty$, ocurre que $p \rightarrow 0$. En un ambiente con población muy cuantiosa infectada de oportunistas errantes y no identificables la estrategia cooperativa sólo puede sobrevivir si el coste personal de las entregas es cero, o sea, sólo entregando bienes sin valor para el cooperador (sin impacto en su eficacia), o interrumpiendo las entregas y con ello el intercambio recíproco (una forma de autarquía atomista en la que en realidad no hay juego, ni tampoco cooperación ni apiñamiento). Efectivamente, los emparejamientos aleatorios con una población lo suficientemente cuantiosa, en la que están mezclados cooperadores y oportunistas, no es posible una diferenciación *a priori*, o la previa discriminación por la vía de demostraciones y señales de fiabilidad es irrealizable, conducen a interacciones que equivalen a juegos anónimos de una sola jugada, y si los jugadores (agentes racionales en este caso) no tienen información acerca del historial cooperativo u oportunista de los demás, la conducta más prudente será reducir al mínimo el nivel de equitatividad (c) para cada expectativa de ganancia dada en el intercambio recíproco (b^*), lo que conduce a una baja cooperación: $\downarrow(c/b^*)$, o incluso a optar por no llevar a cabo transacciones.

El resultado (véase figura 3.4., área no sombreada) es un equilibrio de Nash con cooperación alta si se trata de jugar con conocidos que han cooperado con anterioridad de forma mutuamente satisfactoria (cuadrante superior izquierdo de la matriz)¹⁹¹. Estamos suponiendo que el Jugador A es un cooperador que aplica la regla de reciprocidad siempre que puede. Desde la perspectiva del Jugador A, sin embargo, el pago del juego caerá drásticamente si interacciona con un defector errante (área sombreada derecha en la matriz de la figura 3.4.) En este caso el defector lleva todas las de ganar si la población es numerosa, mientras que el cooperador estará indefenso ante esta situación, pues su posibilidad de represalia en próximas jugadas es reducida o nula. Consecuentemente, el cooperador racional tenderá a una cooperación baja si es capaz de formarse expectativas acerca de la probabilidad de toparse con defectores. La cooperación tenderá a cero si esa expectativa le conduce a ponderar una probabilidad cercana a 1. El proceso lleva implícito una reducción del valor de la variable c , es decir,

(altruismo) o 10 dólares (altruismo y generosidad)» (Zak, *et. al.* 2007: 1). Lógicamente, y aunque los autores no lo especifican, es preciso medir la equitatividad como cuantía relativa a los recursos a disposición de quien hace la entrega (p.ej. su renta). Un altruismo generoso equivale a una forma de altruismo fuerte, mientras que un altruismo de generosidad reducida o basal equivale a un altruismo que podemos llamar pasivo.

¹⁹¹ Obsérvese que esta sería la situación ideal del enclave cooperativo de Axelrod.

una reducción tanto de la cooperación como del nivel de equitatividad¹⁹². En el caso límite, si estima una proporción lo suficientemente elevada de hallarse en un ambiente poblado por abundantes defectores, y/o forma sus expectativas basándose en experiencias pasadas de reiteradas interacciones con defectores, su equitatividad será cero, lo cual equivale a no participar en el juego y retirarse circunstancialmente a la autarquía (fila inferior de la figura 3.4.), como conducta defensiva, hasta que cambien las condiciones del entorno¹⁹³.

Figura 3.4.- Matriz de pagos con valores de (b/c)

		Jugador B			Jugador A
		Cooperación alta	Cooperación baja	Defección	
Cooperación alta		$b'/c = 12/3$	$b'/c = 2/3$	$b'/c = -100/3$	
		$b'/c = 12/3$	$b'/c = 2/3$	$b'/c = -100/3$	
Cooperación baja		$b'/c = 3/1$	$b'/c = (1/3)/1$	$b'/c = -(1/3)/1$	
		$b'/c = 3/1$	$b'/c = (1/3)/1$	$b'/c = -(1/3)/1$	
No se juega		$b'/c = 0$	$b'/c = 0$	$b'/c = 0$	
		$b'/c = 0$	$b'/c = 0$	$b'/c = 0$	

Fuente: elaboración propia.

Nota: obsérvese que cuando $p \rightarrow 0$ la acción más adecuada para el cooperador es aquella que minimice c para los valores posibles de b , puesto que la condición para que su disposición cooperativa pueda evolucionar es $b \cdot p > c$. De este modo, escogerá la conducta de cooperación baja (propia de un altruista prudente), que es en la que menos arriesga *a priori*. Esto se debe a que la incertidumbre ambiental le impide dar verosimilitud a los datos incluidos en la matriz de pagos inicialmente considerada. Su oponente (Jugador B) puede ser un oportunista sofisticado que “esconda un as en la manga” (celdas con fondo sombreado): una defección que le costaría muy cara al cooperador (Jugador A), por lo que, dado el riesgo, la conducta racional por su parte será un nivel de equitatividad bajo que significa la cooperación reducida. Si su historia pasada le permite establecer una distribución de probabilidad suficientemente decantada hacia la expectativa de que su contraparte sea un defector, entonces la elección racional será no participar en el juego, y optar por la autosuficiencia hasta encontrar otro jugador más confiable. En consecuencia, la inexistencia del apiñamiento tiende a mantenerse estable, y resultará menos probable que se formen grupos estables de cooperadores.

Otra forma de explicar este proceso dinámico es la que aparece en el gráfico de la figura 3.5. Como se puede observar, la expresión matemática $p > (c/b)$ puede representarse con dos rectas, que parten del origen y que tienen dos pendientes distintas. Para que la cooperación pueda evolucionar en rondas de interacción reiteradas se ha de cumplir que $\text{tg}(a) > \text{tg}(b)$, o sea que la función que representa la probabilidad asignada

¹⁹² Se trata de lo que Yamagishi, *et. al.* (2005: 299 y ss.) denominan “cooperación cautelosa”, o proceso de cooperación en ausencia de confianza. Esta actitud surge de la incertidumbre asociada al desconocimiento de si se está ante un altruista.

¹⁹³ El resultado que aquí apuntamos es una extensión del análisis de Tullock (1985: 1.079) a un marco en el que no resulta posible seleccionar a los individuos con los que se pretende cooperar. Por eso no funciona, tal y como explica el autor, «la disciplina de los tratos continuos» de la que habló Adam Smith en diversas partes de su obra.

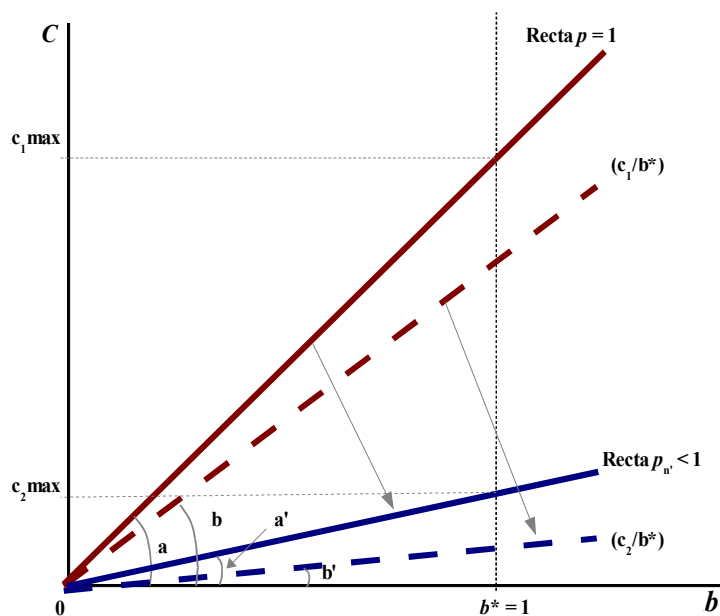
al fenómeno aleatorio “volver a encontrarse con el mismo individuo en una futura interacción” (línea gruesa de color rojo) tenga una pendiente superior a (c/b) , que es el valor de la $\text{tg}(b)$. Si suponemos que b^* tiene un valor constante, por ejemplo $b^* = 1$, la cooperación evolucionará siempre que el valor que adquiriera c_1 sea inferior a $c_1\text{max} = p\text{max} = 1$. Esto quiere decir que el nivel de equitatividad del cooperador A puede situarse entre 0 y $c_1\text{max}$ en el eje de ordenadas. El cooperador puede ser más o menos “espléndido” y a la vez poder sobrevivir como cooperador siempre que $c_1 < c_1\text{max}$ para el nivel máximo de probabilidad $p = 1$. En esta situación de enclave axelrodiano perfecto los individuos pueden comportarse altruistamente en mayor o menor grado, dependiendo de la intensidad de la reciprocidad aplicada, pero, en cualquier caso, preservarán su identidad cooperativa.

El problema surge cuando crece mucho la población n , lo cual, dado que hemos definido p como $p = 1/(n-1)$, lleva a una caída de la pendiente de la recta de probabilidad, que pasa a ser p_n , en línea azul gruesa. Consecuentemente, se desplazará también la recta que representa los valores de c para b^* (línea azul gruesa punteada) para que se mantenga el requisito de que $\text{tg}(a') > \text{tg}(b')$. Como se puede apreciar en el gráfico, ahora la capacidad de elección en cuanto a predisposición de entrega se ve reducida considerablemente, pues $c_2\text{max} < c_1\text{max}$. El cooperador sólo puede entregar b^* siempre y cuando le resulte relativamente poco costoso, pues asume un riesgo bastante mayor de que su contraparte en el juego sea un defector que se aproveche de él no entregándole nada a cambio (o incluso ocasionándole pérdidas adicionales, como en la matriz de la figura 3.4.)

El margen de maniobra para la generosidad, cuando pasamos a un ambiente de extraños que interaccionan sin información acerca de las identidades recíprocas, se reduce con el aumento del tamaño de la población, y eso nos saca del marco conceptual del enclave axelrodiano. Llevando el proceso al límite ($p \approx 0$), el cooperador se verá abocado a minimizar su cooperación, o incluso a suprimirla circunstancialmente, o bien a desarrollar aprendizajes y adaptaciones que le permitan discriminar a los defectores, así como instrumentar socialmente otros sistemas defensivos y disuasorios. El enclave original se ha transformado entonces en un espacio de átomos aislados y sin relación

entre ellos. Este proceso equivaldría, salvando las distancias analíticas, a un descenso abrupto en las relaciones de cooperación dentro de una población dada, con la consecuencia de pasar, en la clasificación de Lorenz (especificada en el apartado 2.2.), del *grupo* o la *superfamilia*, en tendencia descendente de la tasa de generosidad, hacia los estereotipos conductuales de la *sociedad sin amor* o la *multitud anónima*. No es de extrañar que Lorenz atribuyera la capacidad de soportar un elevado coste personal en las relaciones (ya sean de agresión o de altruismo) a los animales capaces de formar grupos estables. Los otros extremos de la escala, el gregarismo pasivo y la indiferencia indiscriminada, no requieren un alto coste personal directo, y no son actitudes enfocadas hacia ningún organismo en particular.

Figura 3.5.- Relación dinámica entre incertidumbre y generosidad



Fuente: elaboración propia.

Los enclaves más frecuentes de la especie humana son apiñamientos del tipo grupo o superfamilia, pero se muestran teóricamente insuficientes para explicar la diversidad de posibilidades de cooperación empírica en agregaciones numerosas pobladas por una gran cantidad de individuos desconocidos entre ellos. Veamos qué mecanismos adicionales pueden favorecer la cooperación con los desconocidos en esta modalidad de sociedades despersonalizadas.

3.3.2.- Reciprocidad indirecta: el valor de la reputación y la justicia comunitaria

El término “reciprocidad indirecta” ha sido presentado por Richard Alexander (1987) para referirse a los cooperadores que no reciben reciprocidad de quien se beneficia de su cooperación, sino de un tercero que conoce lo acontecido y coopera o cooperará con el benefactor inicial. Resulta como si la comunidad en abstracto realizara la reciprocidad, de modo que un cooperador (o un defector) está realizando acciones que serán recompensadas (o castigadas) por alguien que no es su interlocutor directo. La existencia de conductas de reciprocidad indirecta contribuye a reforzar la cooperación en los miembros de la comunidad que son reconocidos por ser cooperadores. En ambientes ricos en información de buena calidad el propio marco institucional permite favorecer a quienes son cooperadores, y excluir de la cooperación a los que no lo son. El concepto clave aquí es el de reputación. Alexander ha visto en esta pauta de reciprocidad el origen de las distinciones de reputación y estatus propias de las sociedades humanas, así como la base de los sistemas morales y jurídicos¹⁹⁴.

El modelo de Alexander tiene un paralelismo coincidente con el pensamiento de Adam Smith, quien se percató de la importancia de la reciprocidad indirecta en la vida social humana más de dos siglos antes, y resaltó el papel instrumental de las instituciones sociales en el comportamiento humano. En efecto, cuanto mayor y más eficaz es la red de información y de cooperación, y mayor la implicación cooperativa del individuo en ella, más relevante resulta la reputación en su comportamiento. La palabra dada de un cooperador determina su comportamiento ante la sociedad, y por eso el avance de la sociedad comercial puede preservar la cooperación a pesar de incrementarse con él la división del trabajo y las interacciones fuera del enclave cooperativo. El sistema de cooperación funciona gracias a los sistemas de compromiso basados en la reputación individual. Dice Smith en las *Lecciones de Jurisprudencia*:

Cuando el comercio es introducido en cualquier país, la probidad y la puntualidad siempre lo acompañan. De todas las naciones de Europa, los holandeses, los más comerciantes, son los que más cumplen con la palabra dada. Los ingleses lo son más que los escoceses, pero mucho menos que los holandeses, y en las partes remotas de este país lo son mucho menos

¹⁹⁴ Un análisis más actual que el de Alexander, dentro de la teoría de juegos, acerca de las relaciones entre reciprocidad indirecta, reputación y normas sociales puede encontrarse igualmente en los trabajos de Ohtsuki e Iwasa (2004), y Brandt y Sigmund (2004).

que en las partes comerciales de él. Esto no ha de ser imputado en absoluto al carácter nacional como algunos pretenden. No existe ninguna razón natural por la que un inglés o un escocés no pudieran ser tan puntuales en cumplir sus compromisos como un holandés. Ello está mucho más relacionado con el propio interés, ese principio general que regula las acciones de todos los hombres, y que les lleva a actuar de cierta forma considerando la ventaja, y está tan implantado en un inglés como en un holandés. Un comerciante teme perder su carácter, y resulta escrupuloso observando cada intercambio. Cuando una persona realiza, pongamos, veinte contratos en un día, no puede ganar mucho tratando de engañar a sus vecinos, ya que la mera apariencia de engaño lo haría perder. Donde la gente negocia poco unos con otros, vemos que está algo más dispuesta a engañar, porque pueden ganar más con una treta ingeniosa que lo que pueden perder por el daño que esto ocasiona en su carácter. (Smith 1978[1762]: 538).

El economista escocés rechaza considerar la cooperación un fenómeno exclusivamente cultural. Su modelo de comportamiento humano pretende explicar universalmente, y por ello engarza en la naturaleza humana la propensión a conservar y señalar el carácter socialmente atractivo por parte de los individuos. Naturalmente, organismos como los humanos, dotados de una buena memoria relacional, están capacitados para reconocer los patrones de comportamiento social que les han funcionado bien en el pasado. Asimismo, están capacitados para reconocer a aquéllos cuyo comportamiento es merecedor de una buena o mala consideración social (ya sea por experiencia propia, inferencia, o simplemente por creer lo que otros expresan acerca de la persona en cuestión). Es indiferente si Smith está flirteando aquí con la psicología utilitaria o está hablando de una psicología de la autoestima basada en la imagen social del sujeto. El caso es que los individuos humanos buscan instintivamente la inclusión social y la pertenencia a grupos, y por ello la buena reputación cooperativa es un valor en sí misma. Ya sea por una vía u otra los individuos tienden a cuidar su buen nombre, sobre todo cuando el contenido de esa información es de acceso público.

Diversos modelos que han tratado de ampliar y mejorar los análisis de Axelrod, Alexander y Smith, han resaltado la relevancia del castigo a los oportunistas y defectores para la supervivencia y estabilidad de la cooperación, lo cual incluye normas y convenciones reguladoras de la reciprocidad más complejas que las recogidas sucintamente en *tft*. Asimismo, se contempla el premio a los cooperadores caracterizados por una generosidad superior a la exigida por las normas sociales. Estos modelos de reciprocidad fuerte (*strong reciprocity*) analizan las consecuencias de la imposición de premios y castigos a los cumplidores e infractores de la norma de

reciprocidad. Esos incentivos, que actúan como refuerzo de la cooperación, no necesariamente han de ser ejecutados por el individuo que haya sido beneficiario de la misma o víctima de conductas anti cooperativas, sino que pueden incluso ser aplicados por otros cooperadores concededores de lo ocurrido¹⁹⁵. La acción de premio o castigo puede estar incorporada a la estrategia básica de reciprocidad enriqueciéndola como pauta de acción altruista, en la medida que ejecutarla suponga un coste no recuperable al individuo que la lleva a cabo (Ones y Putterman 2007: 498). Esta sería la diferencia que establecen Bowles y Gintis (2003: 433) entre lo que ellos llaman *strong reciprocity* frente a la *weak reciprocity* propia de la *tft* de Axelrod¹⁹⁶.

Sin embargo, tal y como apuntan Mises y Wilson en los textos citados en el apartado 2.2., no es necesario asumir que el artífice del premio o el castigo sea un altruista psicológico. Se puede admitir esa conducta como inversión en la pureza de una comunidad cooperativa de la que el *strong reciprocator* es miembro, y de la que se beneficia en el presente y se beneficiará en el futuro tanto él como su prole (un sistema de cooperación multilateral con reciprocidad indirecta). Esta hipótesis conecta con el supuesto teórico de la elección por cálculo utilitario, especialmente si la comunidad actúa como una institución reductora de externalidades y aseguradora de riesgos¹⁹⁷. Además, dichos castigos altruistas administrados por terceros no necesariamente han de ser de un coste elevado para resultar efectivos; pueden consistir simplemente en señalar con el dedo en público, chismorrear sobre una conducta observada, o propiciar la aparición de sentimientos de orgullo o culpabilidad en el

¹⁹⁵ Véanse Fehr y Fischbacher (2003); Fehr y Gächter (2000a); Nowak, *et. al.* (2000); Fehr, *et. al.* (2002); Bowles y Gintis (2004).

¹⁹⁶ En un trabajo más reciente Gintis y Khurana (2008: 310-311) definen la reciprocidad fuerte como: ... la predisposición a cooperar en una tarea de grupo, a responder al comportamiento cooperativo de otros manteniendo o incrementando el propio nivel de cooperación, y a responder al comportamiento no cooperativo de otros castigándolos, incluso soportando un coste personal por ello, e incluso cuando no se pueda esperar razonablemente ganancia personal futura derivada de aplicar ese castigo. Cuando no es posible aplicar otras formas de castigo, el cooperador que aplica reciprocidad fuerte responde a la defección con defección.

¹⁹⁷ Esta hipótesis es contemplada por Bowles y Gintis (1998) en diversos análisis formales. Véanse no obstante las conclusiones de Gary Becker acerca de la caridad social o filantropía como norma moral que sustenta una forma de autoseguro (Becker 1997[1974]: 196-198), y las consecuencias del altruismo familiar en el aseguramiento de sus miembros ante la incertidumbre y en la internalización de externalidades (Becker 1987: cap. 8). Ben-Porath (1980: 7-8) llega a conclusiones similares para el caso del capital relacional en las familias, redes de amigos y empresas, y Buchanan (1965a) y Becker y Murphy (2000: cap. 2) plantean sendos modelos teóricos para explicar la eficiencia de la formación de clubes por individuos con funciones de utilidad homogéneas.

individuo origen del proceso mediante gestos a veces bastante “baratos”, y eficaces¹⁹⁸. Los sentimientos de orgullo y culpabilidad son parte de la arquitectura natural de las emociones sociales humanas, dado que los individuos humanos, al disponer de una memoria desarrollada sobre las conductas propias y ajenas, están capacitados para asociar e identificar a otros individuos con sus comportamientos pasados, y para ser conscientes del valor social de la reputación. En consecuencia, estas emociones condicionan de manera determinante la conducta, favoreciendo tanto la auto-regulación del comportamiento como su acomodación a las relaciones de grupo y control social (Trivers 1978[1971]; Bowles y Gintis 2002; Sigmund, *et. al.* 2002; Simon 1993)¹⁹⁹. Adicionalmente, el individuo que aplica premios y castigos soportando un coste personal por ello también puede beneficiarse de una mejora de su reputación si sus incentivos están alineados con las normas de la comunidad y él puede ser objeto de premios o privilegios de reputación de parte de otros (Mifune, *et. al.* 2010; Henrich y Gil-White 2001). La reciprocidad indirecta confiere una ventaja a los cooperadores que velan por su reputación, y esto permitiría catalogarlos como egoístas a largo plazo que invierten en sí mismos si su motivación es el propio interés descontado al presente (es decir, no salimos de *tft* con tasa de descuento baja).

En este sentido Fehr y Fischbacher (2003: 787) resaltan que una importante cantidad de experimentos realizados con juegos de dilemas sociales acredita que los individuos con una historia de comportamientos generosos hacia los demás tienen una probabilidad significativamente mayor de ser ayudados por otros. Los datos indican que en la conducta cooperativa humana existe una mezcla de motivaciones altruistas y egoístas, puesto que en contextos institucionales en los que la conducta cooperativa, o de refuerzo a la cooperación, se asocia a adquirir una reputación pública, los resultados empíricos corroboran un incremento de los niveles de cooperación en torno al 100%.

¹⁹⁸ Ross (2008: 132) aclara que los individuos humanos están biológicamente adaptados para ser muy sensibles a pequeños premios (sonrisas, miradas de complicidad, saludos etc.) y castigos (negar la conversación, esquivar la mirada, realizar una desaprobación, etc.) relacionales que no conllevan un coste significativo para quien los imparte, pero que sí pueden producir un impacto mucho mayor en el individuo que los recibe. Según Ross algunos autores han exagerado el coste individual de esos premios y castigos, haciendo de quien los ejecuta un altruista fuerte que no necesariamente lo es.

¹⁹⁹ Las normas sociales y morales consideran execrables los comportamientos antisociales y las personalidades excesivamente narcisistas porque la aversión a la antisocialidad tiene carácter epigenético (Sommer, *et. al.*, 2010). La psiquiatría moderna, por su parte, incluye entre las patologías clínicas graves las carencias en los mecanismos neurofisiológicos que regulan la producción de emociones sociales. Véase Lykken (1995) sobre trastornos de personalidad antisocial.

Parece claro que la motivación es sensible a las posibilidades de mejorar la reputación personal y, obviamente, la reputación es un activo valioso cuando existe observación individualizada (*peer-pressure*) y la información acerca de la conducta puede fluir con facilidad hacia personas con quienes exista una probabilidad significativa de interactuar en el futuro. Este es un asunto relevante para la economía, pues conecta con las teorías de los costes de transacción y de los juegos cooperativos:

Imagine un mundo donde cualquiera pudiera ser conocido con una completa seguridad, donde la gente no incurre en ningún coste de cálculo, recuerdo y reconocimiento de otras personas; y donde cada transacción -ya sea aquí y ahora, o a través del tiempo, el espacio u otras inciertas contingencias- puede realizarse sin costes (de transacción). En ese mundo las personas podrían hacer negocios entre ellas simplemente con el respaldo de su palabra. No habría necesidad del dinero. (Dasgupta 2007: 54; paréntesis añadido).

3.3.2.1.- Reputación y sistema institucional

La reputación es un tema que requiere una análisis más extenso, pues los modelos evolucionistas ensalzan su papel en la explicación del comportamiento cooperativo dentro de interacciones multipersonales más complejas (y realistas) que las sencillas díadas de Axelrod. Cuando hablamos de reputación podemos referirnos a una “buena” o a una “mala” reputación. Generalmente la primera viene asociada a una serie de etiquetas morales y culturales humanas, tales como la honestidad, la fidelidad, la generosidad, el altruismo y en general un talante benevolente, generoso y cooperativo. La mala reputación vendría a ser la negación de la buena, es decir, su contrario: el egoísmo predatorio u oportunista en sus manifestaciones conductuales. Los especialistas aún no han llegado a un acuerdo sobre el concepto de la buena (mala) reputación y cómo introducirlo en los modelos evolutivos de reciprocidad indirecta (Fehr y Fischbacher 2003: 789). La pregunta clave parece ser ¿qué tipo de comportamientos “bondadosos” permiten la evolución de la cooperación por medio de la reciprocidad indirecta?

Según el modelo creado por Ohtsuki e Iwasa (2004) los juicios morales sobre el comportamiento observado son realizados según la siguiente regulación²⁰⁰:

²⁰⁰ Estas reglas no necesariamente son productos de la razón humana, sino el resultado de procesos coevolutivos (Boyd y Richerson 2008).

- a) La cooperación con buenas personas se considera buena, mientras que la defección en contra de esas mismas buenas personas es una conducta mala y reprochable;
- b) la defección con malas personas es un comportamiento bueno, pues supone de hecho un castigo; del mismo modo la cooperación con malas personas puede considerarse un premio, y por lo tanto un incentivo inapropiado para la evolución de la cooperación.

Sin embargo, queda la incógnita de qué es una buena (mala) persona. El sentido común indica que alguien es bueno si es un cooperador, y es malo si actúa como defector o *free-rider*. Pero, ¿traicionar a un traidor es una conducta bondadosa o malvada? Una posible salida de esta disyuntiva puede ser alinear la corrección de la conducta con aquello que establecen las normas sociales, considerando buenos los actos que cumplen las normas, y viceversa. En este sentido, castigar con defección a quienes violan las normas sería algo bueno (Fehr y Fischbacher 2003: 786). Esta regla está alineada con la preservación de la estabilidad y orden social; se trata por tanto de una estrategia conformista de ordenación social.

Así pues, la bondad o maldad de los actos e intenciones es la variable explicativa de la reputación socialmente atribuida a un sujeto, que siempre consiste en un bien relacional con valor *relativo* en la interacción social. Ohtsuki e Iwasa (2004: 117) citan un trabajo realizado por Takahashi y Mashima en el que se concluye que la reputación del participante en una interacción social está determinada por:

- a) Su acción, cooperativa o no (información de primer orden), y
- b) la reputación (cooperativa o no) de su oponente (información de segundo orden).

Siguiendo esta clasificación, el cooperar con cooperadores haría ganar reputación al agente, mientras que quien traiciona o se aprovecha de los defectores también recibiría

el premio de la reputación. Pero queda sin contestar un tema que consideramos relevante, y es el de que todo esto está orientado por el sentido moral de las normas sociales, y las normas sociales no comprometen por igual a todos los colectivos. Es más, cada grupo tiene normas propias, y en las acciones inter-grupales pueden aparecer conflictos en cuanto a la calificación moral de las acciones. De este modo, ¿aporta reputación cooperar con un miembro de otro grupo que coopera con los suyos de acuerdo a sus normas, pero que está, como grupo, enfrentado al propio? Da la impresión de que el valor moral de la reputación se circunscribe al contorno normativo-cultural de cada grupo, y que las alianzas cooperativas con grupos ajenos tienen valoración moral favorable si existe una semejanza de rasgos normativo-culturales. Esto apoya la tesis de un relativismo moral localista y grupalista dependiente de la evolución cultural, asentado en una estructura instintiva (*hard wired*) de la justicia que es universal (Binmore 2006a: 23; Elster 1994; Hauser 2006; Greene 2009). Tal estructura permite la adaptación a las convenciones reguladoras de la ordenación del comportamiento social, estableciendo puntos focales que evitan la problemática de los equilibrios múltiples²⁰¹.

Russell Hardin (2002: 126-127) explica dos ejemplos empíricos ilustrativos al respecto de la reputación y su papel en las transacciones económicas. El primero es el del valor de la reputación en el sistema financiero informal que existió en la comunidad de inmigrantes cubanos huidos de su país a Miami tras llegar al poder Fidel Castro. La reputación empresarial que muchos de esos emigrantes fraguaron en Cuba fue la llave que permitió abrir la puerta de la confianza en el exilio, y les permitió acceder a préstamos, sin garantías, de entre 10.000 y 30.000 dólares de aquella época para crear nuevos negocios en EE. UU. El segundo ejemplo que cita Hardin es el del sistema originario de microcréditos creado por Muhammad Yunus en Bangladesh. En este caso,

²⁰¹ Darwin (1972[1871]: cap. 5) concebía la moralidad reguladora de la conducta humana como un bien de propiedad grupal, y por lo tanto circunscrita al ámbito de grupo. De este modo, las reputaciones se atribuyen en función de la pertenencia a un colectivo y la coherencia o no de la conducta individual con los intereses de ese colectivo. Esto da sentido a las expresiones populares de “amigo” o “enemigo” público. Así, un individuo que coopere con un sagaz oportunista que se aprovecha de los miembros del grupo verá perjudicada su reputación, mientras que si ese sagaz oportunista es una especie de Robin Hood que se dedica a robar a los miembros de un grupo rival, entonces su acción será premiada con una buena reputación. En el fondo se trata de una cuestión de afinidades simpatéticas e identidad grupal que supo captar acertadamente Bentham (2000[1781]: VI, xxvii):

La simpatía hacia tu amigo produce el nacimiento de *tu* antipatía contra todos aquellos que *le* resultan antipáticos, así como la simpatía por todos aquellos que son objetos de *su* simpatía (...) La antipatía contra tu enemigo genera simpatías *hacia tí* por parte de aquellos que son objetos de su antipatía, así como de antipatía por parte de los que son objetos de su simpatía. (Cursivas en el original).

la confianza se asocia a la buena reputación no sólo del prestatario, sino del conjunto de su grupo relacional directo de familiares y amigos. De este modo, los grupos en que uno de sus miembros no ha devuelto el préstamo recibido en las condiciones pactadas son penalizados con la pérdida de reputación como si se tratara de un conjunto integrado y con una identidad social única. Este segundo caso contempla las dos versiones de la reciprocidad indirecta descritas: quien tiene una historia de cooperación resulta premiado, y además existen incentivos para que los miembros de su grupo vigilen que la reputación conjunta sea preservada, aunque no necesiten un préstamo en ese momento.

Conviene resaltar que el sistema de la reciprocidad indirecta sólo puede funcionar cuando existe información barata, accesible y fidedigna sobre la reputación. Cuando esto es así, la reputación tiene valor; de lo contrario se trata de un concepto no contrastable (Ohtsuki e Iwasa 2004: 107-108; Nowak 2006a: 1.561). Nowak, *et. al.* (2000: 1.773) encuentran una significativa correlación entre la evolución de la equidad (*fairness*) y la evolución de la cooperación, en el sentido de que ambas están asociadas a la reputación. La publicidad de la información y la eficacia comunicativa son aspectos clave, y prueba de ello es que sólo en algunas especies cognitivamente evolucionadas se han encontrado sencillos sistemas de reciprocidad indirecta, aunque ninguno tan sofisticado como el de la especie humana. De hecho, el sistema de reciprocidad indirecta basado en la reputación se da en su forma más desarrollada sólo en las sociedades humanas, dado que requiere la disposición de capacidades cognitivas como la memoria, la vigilancia y comprensión de las intenciones ajenas, el lenguaje y el manejo de estereotipos y categorías sociales (Nowak 2006a: 1.561). No es de extrañar que, siendo esto así, una de las formas de competencia intra e inter grupal dentro de las sociedades humanas consista en la lucha por el control y la manipulación de los canales públicos de información, y que uno de los primeros asaltos que llevan a cabo quienes pretenden actuar totalitariamente (esto es, anti-cooperativamente) sea el control y la tergiversación de la información sobre las reputaciones, a través de la propaganda y la intoxicación informativa.

Siguiendo este modelo, es preciso pues que la reputación sea señalizable con fidelidad, y que la señalización sea un marcador de conocimiento público y fiable.

Aunque sostenemos, con Dawkins (1993[1976]: 245), que la especie humana dispone de una bien desarrollada memoria y capacidad de reconocimiento de otros individuos, sin embargo, estas capacidades están limitadas al ámbito extensivo de un grupo de tamaño limitado al número de Dunbar (o sea, 120 – 150 individuos)²⁰². La mayor parte de las redes de interacción informal en las que se actúa hoy día, y el tamaño de las unidades organizativas formales de las civilizaciones post-neolíticas, superan y han superado frecuente y ampliamente la barrera de Dunbar. Un desarrollo neurológico epigenético, con incidencia cultural e institucional, e igualmente con *feed-back* coevolutivo, es la capacidad de cálculo aplicada a la medida de la equitatividad en las transacciones, y su reflejo material e institucional: el dinero. Parafraseando a Dawkins (1993[1976]: 245):

De especial interés son los “engañadores sutiles” que parecen estar pagando un favor recibido pero que, sin cesar, devuelven menos de lo que reciben. Es aun posible que el abultado cerebro humano y su predisposición a razonar matemáticamente haya desarrollado un mecanismo de engaño más tortuoso, así como una detección más penetrante del engaño cometido por otros. El dinero constituye un signo formal del altruismo recíproco retardado.

Desde una perspectiva evolutiva puede comprenderse la reputación abstracta como un antecedente del dinero. Si bien la reciprocidad directa se asemeja a un trueque bilateral, en el caso de la reciprocidad indirecta el cooperador acumula derechos a recibir bienes de los demás. O sea, quien posee buena reputación es un acreedor del resto de la sociedad, y un desarrollo institucional que pudo evolucionar en las sociedades primitivas para reconocer esos derechos bien pudo ser lo que hoy llamamos dinero. Efectivamente, existen pruebas de que las conductas de los humanos modernos se remontan a unos 90.000 – 100.000 años atrás y, si bien las primeras sociedades con una división del trabajo diferencialmente compleja surgieron en Mesopotamia hace 5.000 años, ya antes de la revolución neolítica se daban tratos comerciales dentro de las tribus y entre ellas, utilizando medios de cambio estandarizados como conchas, piedras, u otros objetos especiales (Richerson y Boyd 2008: 108).

²⁰² Robin Dunbar (1996) propone que los humanos han vivido durante la mayor parte de la historia cooperando dentro de grupos estables de entre 50 y 150 individuos. La teoría de Dunbar especifica que el tamaño del neocórtex humano moderno limita la posibilidad de desarrollar interacciones de cooperación fuerte en poblaciones por encima de 150 individuos.

Esta consideración del dinero como medio de pago que facilita el comercio no resulta sorprendente para el economista. Pero quizás resulte menos conocido el papel que representaron ciertos objetos como grandes piedras en forma de rueda que han sustentado primitivos sistemas monetarios, como el de los habitantes de la Isla de Uap en Micronesia (véase Friedman 1992: cap. 1). Como relata Friedman, dichas voluminosas y pesadas ruedas de piedra tenían un agujero central que se usaba para moverlas, aunque algunas familias muy “ricas” no necesitaron usarlas nunca como medio de pago, y su “dinero” permanecía depositado en el fondo del mar como señal inmóvil del prestigio de un linaje. Es de suponer que estas familias eran buenas “ahorradoras”, y “producían” más de lo que “consumían”. Cuando hablamos aquí de ahorro estamos aludiendo a individuos que han ayudado cooperativamente a muchos, o que han ayudado mucho a unos pocos. Una forma de reconocer esa deuda en términos de reciprocidad indirecta por el resto de la sociedad es la señalización mediante la posesión de objetos que convencionalmente acreditan que la comunidad está en deuda con su poseedor. Esto significa que el poseedor es una persona con buena reputación, y que esa reputación es reconocida por el resto de miembros de su grupo cultural²⁰³.

Obviamente, las sociedades actuales son mucho más complejas que las tribus a las que alude Friedman y, si bien en las comunidades arcaicas de pequeño tamaño todos sus miembros tenían información sobre la legitimidad de la posesión de ese dinero, en las sociedades modernas existen con frecuencia elevados costes de transacción para reconocer si el dinero que uno posee ha sido obtenido cooperando o, por el contrario, engañando y explotando a los demás²⁰⁴. En las sociedades complejas el dinero tiene la principal función de ser un medio de intercambio, pero no se puede pasar por alto su función señalizadora de la reputación, especialmente en aquellos casos en que los costes de transacción relacionados con la verificación de esa reputación son moderados o

²⁰³ La psicología moderna reconoce que la motivación de acumular dinero no necesariamente es una motivación intrínseca, sino que suele ser instrumento de otras motivaciones de orden superior. Algunas de ellas son de claro contenido prosocial, y aportan bienestar psíquico al individuo en la interacción social. Véase Srivastava, *et. al.* (2001).

²⁰⁴ Como propone Scharpf (1990) un prerrequisito básico para el éxito de relaciones sociales cooperativas, en el caso de animales con racionalidad limitada como los humanos, es la existencia de instituciones proveedoras de un conocimiento público acerca de los pagos económicos del juego, la historia de los jugadores, así como de sus estrategias e intenciones en situaciones sociales determinadas. En definitiva, se trataría de disponer de información fidedigna sobre las reputaciones.

irrelevantes²⁰⁵. No sorprende pues que los modelos manejados concluyan con que existe una relación matemática entre la propagación de la cooperación vía reciprocidad indirecta y el conocimiento que se tiene acerca de la reputación de los demás (Nowak y Sigmund 2005; Milinski, *et. al.* 2002; Panchanathan y Boyd 2004). Así, la relación que existe entre el coste personal que soporta un individuo respecto al beneficio de la reciprocidad indirecta con la que recompensa a quien considera merecedor de su reciprocidad fuerte (c/b), ha de guardar una relación con la probabilidad (q) de conocer la reputación del beneficiario que explica la siguiente desigualdad²⁰⁶:

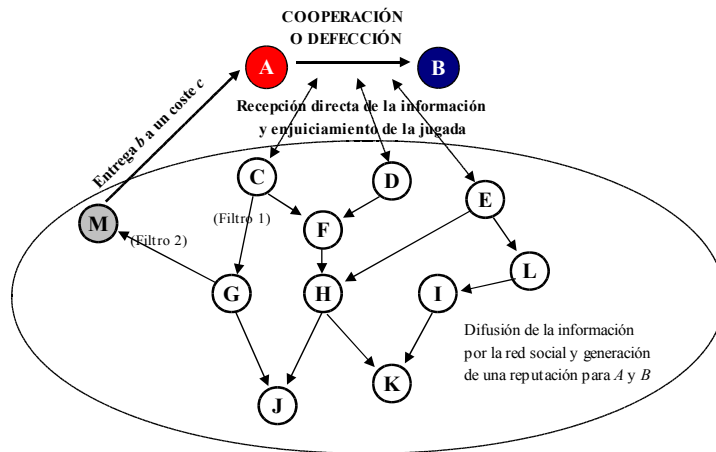
$$q > c/b$$

La reciprocidad indirecta permite la evolución de la cooperación si la probabilidad de disponer de una buena información acerca de la reputación ajena es superior a la relación coste/beneficio aludida (véase figura 3.6.)

²⁰⁵ En las sociedades modernas la propiedad de riquezas tiene una valoración en términos monetarios, y el dinero es riqueza en sí mismo. Quien posee riquezas lo da a conocer haciendo ostentación de ellas, porque la riqueza es una señal de estatus dentro de la estructura social humana (Veblen 2002 [1899]). El estatus, de otra parte, favorece la seguridad y la replicación biológica del individuo. El dinero es una señal generalmente más fácil de captar por desconocidos que otras habilidades o facultades valiosas (como la inteligencia, la honestidad o la pertenencia a una saga familiar poderosa).

²⁰⁶ La demostración del teorema se encuentra en Nowak (2006a: 1.561).

Figura 3.6.- Esquema de la reciprocidad indirecta



Los individuos *A* y *B* intervienen en una interacción cuyo resultado es cooperación o defección por parte de *A* y/o *B*. Los individuos *C*, *D* y *E* observan directamente el juego y efectúan una valoración emocional de los comportamientos, y difunden la información por el resto de la red. La información llega filtrada a *M*, quien realiza una valoración en función de la credibilidad de la información recibida, con la que atribuye una determinada reputación a *A*. La calidad de la información que recibe *M* depende de la cantidad de nodos por los que haya circulado la información acerca de los jugadores, así como la afinidad genética y cultural que exista entre *M* y los nodos intermedios. La probabilidad (*q*) de conocer la reputación de *A* es una función de la cantidad de nodos intermedios (*n*) así como del nivel de afinidad (ζ) existente entre *M* y esos individuos:

$$q_{MA} = I(\zeta_1 \cdot \zeta_2 \cdot \zeta_3 \dots \zeta_n)/n ; 0 < \zeta < 1$$

Siendo *I* la información en cuestión acerca de la reputación de *A*. El individuo *M* actuará premiando o castigando al individuo *A* si $q_{MA} > (c/b)$.

En nuestro ejemplo, donde existen 2 nodos intermedios: $q_{MA} = I(\zeta_1 \cdot \zeta_2)/2$

Fuente: elaboración propia inspirada en Nowak (2006a: 1.561) y Ohtsuki e Iwasa (2004: 110).

3.4.- Limitaciones del enfoque de la reciprocidad racional para explicar la evolución de la cooperación

La evolución de la cooperación en los experimentos de Axelrod (1996) requiere el cumplimiento de unos supuestos ciertamente restrictivos, díganse:

- a) Interacciones bilaterales repetidas indefinidamente,
- b) en las que la “sombra del futuro” influye de forma decisiva en la acción presente (en los cálculos utilitarios de los agentes económicos, si asumimos que los

jugadores son humanos), y

- c) los individuos tienen acceso a una buena y barata información acerca del historial de la contraparte en las transacciones, o sea, escasos o nulos costes de transacción.

El modelo victorioso de los torneos de estrategia organizados por Axelrod aporta información teórica que permite explicar cómo pudo fructificar la cooperación en pequeños grupos tribales prehistóricos de 50-150 individuos, con una limitada movilidad individual inter-grupos (Dunbar 1992 y 1996; Tooby, *et. al.* 2006). Igualmente puede arrojar luz sobre la naturaleza de las comunidades de reducida dimensión existentes en la actualidad, cuya cultura impone severas limitaciones a las relaciones con *outsiders* (véase Shenker 1986). Estos grupos, bastante comunes en otras especies de mamíferos sociales, se caracterizan por su alta “viscosidad”, concepto que alude a «cualquier tendencia de los individuos para continuar viviendo juntos en el lugar donde nacieron» (Dawkins 1993[1976]: 282). La viscosidad contribuye a generar semejanzas genotípicas y fenotípicas en las poblaciones locales a través de la evolución por deriva (genética y cultural). Este proceso determina que los individuos tengan una alta probabilidad de interactuar con otros que son muy parecidos conductualmente hablando (esto es, con estrategias que son buenas copias). Como ya hemos explicado, se trata de uno de los requisitos para que una estrategia eficaz, como *tft*, sea evolutivamente estable. Resulta comprensible la relación entre la viscosidad de una población y el cumplimiento de los tres requisitos básicos de Axelrod para la evolución de la cooperación. El modelo recoge marcos relacionales estables e informativamente ricos, además de llevar implícita la interdependencia de los componentes del grupo debido al prerrequisito de viscosidad.

Una de las críticas más contundentes que se puede hacer a esta teoría de la cooperación consiste en que las estrategias que compiten no evolucionan, lo cual es una negación del propio significado de la evolución en los sistemas vivos²⁰⁷. La complejidad

²⁰⁷ Para auténticos juegos evolutivos véanse los trabajos posteriores del propio Axelrod (1997: caps. 1 y 2), y otros realizados por Nowak (2006b); Nowak y Sigmund (2004), Kraines y Kraines (1995), Nowak, *et. al.* (1995), Sigmund (1993), Nowak y May (1992), Levy (1992), Langton (1989).

real de la evolución es mucho mayor que la recogida en el trabajo de Axelrod, quien tenía una visión de la evolución darwinista limitada a la modificación incremental. Así, un cambio aleatorio en el ambiente, que puede consistir, por ejemplo, en una incursión de defectores *outsiders* en el grupo estable de cooperadores; o un aumento significativo del tamaño del propio grupo de cooperadores; la súbita necesidad de interactuar con los extraños “de afuera” ante una catástrofe ecológica; la reducción drástica de la esperanza de vida de los individuos o del tiempo esperado de existencia del grupo, o la transcripción errónea de la estrategia en su ejecución real (una forma de mutación), pueden dar al traste con el éxito evolutivo de *tft* y provocar la selección de otras estrategias dominantes no cooperativas (Nowak y Sigmund 1992 y 1993a).

Pero dejando de lado el cambio y las innovaciones de comportamiento, aun en el plano estático, Axelrod es consciente de la precariedad de la cooperación basada en la reciprocidad directa, y, no en vano, dedica buena parte de su libro (los capítulos 6, 7 y 8) a explicar fórmulas “intervencionistas” que permitan reforzar la cooperación cuando no se cumplen los requisitos básicos mencionados²⁰⁸. En efecto, los incentivos a la cooperación del tipo *tft*, se dan en unas condiciones limitadas dentro de las sociedades humanas complejas, especialmente en las grandes sociedades de los países industrializados desde al menos los dos últimos siglos (Eibl-Eibesfeldt 1996). Como se sabe, *tft* comienza cooperando, pero en lo que no se suele reparar es en que en encuentros esporádicos con individuos desconocidos, a los que se presume que no se volverá a ver, no existe justificación racional para cooperar. Así, North (1993: 79) reconoce que en el mundo de intercambios impersonales, con información imperfecta y costosa, donde interactúan muchos individuos que no se conocen en jugadas que en principio no han de repetirse, la estrategia *tft* no será evolutivamente estable a no ser que venga reforzada por arreglos institucionales *ad hoc*, lo cual nos devuelve en parte a la microeconomía de la no-cooperación y sus problemas asociados²⁰⁹. Pero sería erróneo

²⁰⁸ En esencia lo que el autor está planteando es una manipulación racional del ambiente que propicie la estabilidad relacional, el reconocimiento de los individuos en función de los resultados de interacciones precedentes, y la selección según tales resultados. Para ello recomienda diversas posibilidades complementarias y no excluyentes de intervenir en el marco institucional: aumentar el peso del futuro en las decisiones individuales, interferir en el sistema de incentivos a través de premios y castigos, y «enseñar valores, hechos y destrezas» a modo de una ética de la reciprocidad.

²⁰⁹ Cabe precisar que North ha introducido en modelos posteriores el factor “creencias” como soporte cultural de las preferencias e instituciones que permiten una cooperación generalizada. Por ejemplo North (2005).

extrapolar sin más estas tendencias conductuales no-cooperativas a las empresas y otras organizaciones también de las sociedades modernas, dado que en la mayor parte de ellas se dan los requisitos básicos para la cooperación que señala Axelrod, al menos en cierto grado²¹⁰. Sin embargo, la mayoría de enfoques de economía de la organización tratan las relaciones internas entre individuos de manera idéntica a las interacciones con extraños. Esto puede deberse a que tanto los mercados como las organizaciones formales que teoriza la microeconomía neoclásica sean ámbitos de anonimato, pobres en información y comunicación, o tengan una estructura poco concentrada, como para que proliferen relaciones personalizadas típicas de un ambiente de interacción densa (véase figura 3.7.)

Figura 3.7.- Niveles comparativos en la densidad relacional

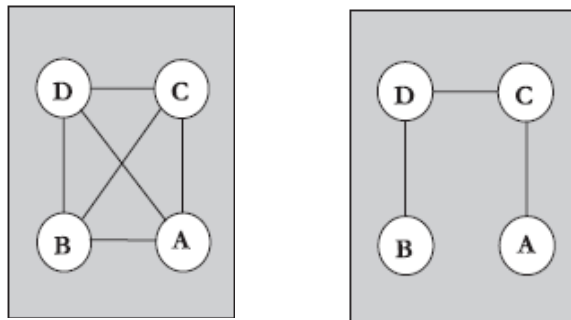


Fig 7.1.- Grupo de alta densidad ($\Delta = 1$) Fig 7.2.- Grupo de baja densidad ($\Delta = 0,5$)

Los círculos indican individuos que integran un grupo (A, B, C, D), y las líneas indican flujos relacionales de comunicación y reciprocidad bilateral. La densidad relacional puede medirse como el cociente del número de relaciones existentes en un grupo y la suma total de relaciones potenciales en ese mismo grupo. Formalmente se expresa mediante la siguiente ecuación: $\Delta = 2R / I(I-1)$, donde R es el número de relaciones existentes, e I es el número de individuos que integran el grupo. El valor de Δ está siempre situado entre 0 y 1, en un *continuum* que va desde la inexistencia de flujos relacionales ($\Delta = 0$; lo que equivale a la inexistencia práctica de grupo), hasta la presencia de abundantes flujos en los que cada miembro tiene relaciones con todos los demás ($\Delta = 1$; que equivale a un grupo completamente integrado).

Fuente: Levine y Kurzban (2006: 177).

²¹⁰ Durante la mayor parte de la historia de la especie, los humanos han evolucionado en ambientes y poblaciones cognitivamente limitados por el número de Dunbar. Nunca antes el individuo humano había realizado interacciones esporádicas por encima de ese límite cognitivo y emocional con la alta frecuencia con que se llevan a cabo en las sociedades contemporáneas (Eibl-Eibesfeldt 1996: 16). Ello genera incertidumbre y desconfianza, e importantes dosis de estrés. No obstante, ese incremento de la desconfianza y la incertidumbre en las interacciones no necesariamente ha de reducir la cooperación a cero, ya que de ser así la civilización actual sería inviable a pesar de las salvaguardas *à la* Williamson y la presencia del Leviatán (North 1984: 61).

Así, Greif (2004: 9) da a entender que muchas relaciones de intercambio, incluso dentro de una organización con estructura formal, conllevan implícitamente una necesaria «separación temporal y/o geográfica entre el *quid* y el *quo* intercambiados», lo que socava la reciprocidad por el lado de la información. Otras razones podrían apuntar a factores de tipo motivacional, como el conflicto inherente a las relaciones de producción capitalistas sobre el que trata la tradición marxista, o cualquier otro tipo de conflicto de intereses de otra índole a parte del ya implícito en el estereotipo conductista de la relación de agencia o el *free rider* en la acción colectiva con bienes públicos (como p.ej. heterogeneidad de culturas; creencias excluyentes; normas informales internas perniciosas o discriminatorias hacia ciertos miembros del grupo, o un ambiente psico-social de crispación y desconfianza)²¹¹. Chester Barnard (1968[1938]: 162-164) insistió en la importancia para la cooperación efectiva en las organizaciones de que tanto las normas internas, como los sistemas de recompensa, fueran percibidos como justos y legítimos por sus miembros, una idea que también han recogido Arrow (1974: 71 y ss.) y el propio North (1984: cap. 5).

En el marco de relaciones de cooperación sin Estado que estamos analizando han sido señalados diversos mecanismos de control y disuasión del oportunismo utilizados en épocas históricas en las que la cooperación no contaba con el refuerzo del sistema jurídico-político. Se trata de la diversidad de modelos de organización ante el problema de que puedan existir defectores difícilmente identificables dentro del grupo. El problema fundamental del intercambio ha sido atajado con ingeniosos sistemas descentralizados, algunos de ellos todavía hoy utilizados, como el intercambio o toma de garantías, pagos diferidos, defensa multilateral de la reputación étnica o religiosa, exclusividad de relaciones dentro de redes estables con castigo multilateral al oportunista, búsqueda de la homogeneidad cultural, sistemas de responsabilidad solidaria comunitaria, etc. (véanse Greif 2004; Ostrom 1998). Se trata de generar expectativas fundadas (“compromisos creíbles”, en palabras de Williamson 1983) para propiciar la inusual confianza entre individuos que no pertenecen a una misma familia biológica o cultural-territorial y, más que establecer relaciones duraderas de proximidad

²¹¹ En este aspecto Williamson ha destacado la relevancia económica de cuestiones tales como las preferencias por el ambiente, la “atmósfera” y la dignidad en el diseño de las organizaciones humanas (Williamson 1989: 272-275; y 1991: 55-58).

(al estilo del apiñamiento que describe Axelrod), los incentivos para el autocumplimiento de los compromisos asumidos tienen que ver con la información y la posibilidad de aplicar un castigo lo suficientemente creíble y dañino como para disuadir de la tentación a los oportunistas (véase Akerlof y Yellen 1986). Douglass North (1993: 61 y 78) añade además que las creencias comunes y compromisos ideológicos dentro de poblaciones de acceso restringido han permitido, en algunos tiempos y lugares, crear alianzas para el intercambio dentro de grupos dispersos geográficamente, pero dotados de un notable conocimiento mutuo y estabilidad relacional (una alta viscosidad social), además de un autocontrol del oportunismo mediante la fidelidad identitaria a ciertos códigos de conducta estrictos y exclusivos²¹².

Pero el principal problema al que nos enfrentamos, admitida la necesidad de compromisos creíbles para reforzar la cooperación, tiene que ver con dos cuestiones que dependen en gran medida del tamaño de la red de interacción, y que se relacionan tanto con las características de los individuos que la integran, como con las características del sistema institucional necesario para disuadir y castigar el oportunismo, y coordinar la cooperación:

- a) Los grandes grupos y las grandes redes tienen una mayor probabilidad de que sus individuos tengan pautas de comportamiento o preferencias heterogéneas; es decir, es más probable que al expandirse la frontera y la gama de relaciones posibles se mezclen individuos con tendencia a la cooperación recíproca e individuos con tendencia a la defección, y que sus interacciones sean juegos de una sola tirada. Asimismo, conforme la heterogeneidad aumenta es menos probable encontrar objetivos y pautas de acción comunes, lo que dificulta que aparezcan procesos de interdependencia cooperativa;

²¹² Kandori (1992) muestra cómo el refuerzo de la cohesión grupal a través de la interiorización de normas y la difusión de la confianza es relevante para explicar la evolución de la cooperación. En este sentido, Tirole (1996) ha presentado un modelo sobre la corrupción en las organizaciones en el que muestra cómo la pertenencia a grupos sociales (organizaciones) de alta reputación genera incentivos a cooperar, dado que la mera pertenencia al grupo (que se supone lleva implícita el cumplimiento de las normas internas) aporta una renta que se perdería si el individuo es expulsado por comportarse oportunistamente. Amartya Sen ha analizado en varios de sus trabajos la relación existente entre el compromiso con los objetivos de otros jugadores pertenecientes a un determinado grupo (con el cual el individuo se identifica) y el éxito de la cooperación (véanse Sen 1985 y 1986).

- b) dicho esto, los mecanismos de información y ejecución dirigidos al control del oportunismo y su castigo, y a coordinar las relaciones de cooperación mediante reglas grupales, tenderán a ser más costosos y complejos, además de ineficaces, especialmente si admitimos que con el aumento del número de individuos tiende a disminuir la densidad relacional, y por lo tanto la capacidad multilateral efectiva de controlar el cumplimiento de la cooperación y la información acerca de las reputaciones²¹³. Si el control y la coordinación resultan costosos o imposibles, entonces el problema de la acción colectiva caerá por la pendiente de la provisión conjunta de bienes públicos en un ambiente típico de *free-riders*, es decir, la tragedia comunal clásica (Olson 1971[1965]; Hardin 1968).

La estructura argumental que exponemos a continuación se asemeja a la que plantean Bandiera, *et. al.* (2005), así como Hechter y Kanazawa (1993). En esencia se trata de contemplar dos propiedades que tienen que ver con la naturaleza de los individuos que comparten un sistema, así como de las instituciones del propio sistema. Los autores ofrecen una teoría del orden social que no requiere necesariamente la figura de un tercer agente especializado en reforzar las reglas de cooperación (como pueda ser el Estado, o el gobierno corporativo de las organizaciones)²¹⁴. En este sentido se trata de un análisis del orden espontáneo. Dichas propiedades son las siguientes:

- a) *Interdependencia/complementariedad*: cuanto más interdependientes sean las funciones de utilidad de los individuos, menor será el coste relativo del control mutuo requerido para conseguir elevados niveles de cooperación. Esto tiene que ver con varios aspectos, tales como las externalidades positivas cruzadas derivadas de una mayor homogeneidad en las preferencias, o de utilidades individuales correlacionadas positivamente. Asimismo, tiene que ver con el alineamiento de incentivos al desempeño cooperativo por otras vías instrumentales, como las instituciones gubernamentales o el contrato privado.

²¹³ Bowles, *et. al.* (1997: 7) equiparan el grado de densidad relacional que existe en una población al concepto de capital social, que consideran como el grado en que los individuos pueden supervisar el comportamiento de los demás integrantes de la población.

²¹⁴ Un análisis microeconómico de referencia sobre el marco contractual de la cooperación entre maximizadores egoístas, con agente especializado en reforzar la cooperación con incentivos, es el trabajo de Alchian y Demsetz (1994[1972]).

- b) *Visibilidad*: cuanto más elevadas sean las capacidades cognitivas de los individuos, y superior el nivel de publicidad de intenciones y comportamientos de los mismos en el desarrollo de sus acciones, menores serán los costes de vigilancia y eventual sanción por la vía de la reciprocidad negativa a los no-cooperadores.

Según Hechter y Kanazawa (1993: 461) la interdependencia y la visibilidad son dos aspectos explicativos del surgimiento de relaciones sociales cooperativas sin necesidad de Estado o figuras institucionales semejantes, si bien no son los únicos. En cualquier caso, el orden cooperativo de los pequeños grupos sociales se explica por ambas cualidades. El enclave axelrodiano es un grupo con elevado nivel de viscosidad relacional, en el que hay altos niveles de interdependencia y visibilidad. Explicamos la problemática asociada a la ausencia de una o ambas cualidades estructurales en sociedades más complejas.

3.4.1.- El problema de la interdependencia/complementariedad

Dos o más individuos son interdependientes si existe una posibilidad de interacción entre ellos que resulte óptimo-paretiana respecto a la ausencia de relación alguna. En los grupos de individuos interdependientes se da complementariedad en el sentido de que pueden existir sinergias derivadas de la acción en el contexto grupal. Un factor explicativo del gregarismo pro-grupal es la homogeneidad, que posibilita objetivos y procedimientos compartidos en la interacción. El caso más evidente de interdependencia se produce cuando los individuos no pueden sobrevivir o encuentran graves dificultades para ello fuera del agrupamiento.

A veces se entiende de forma distorsionada la capacidad individual de los animales sociales para sobrevivir autónomamente fuera de grupos cooperativos. Sin embargo, es sabido que los individuos de estas especies reducen drásticamente sus posibilidades de supervivencia fuera de las agrupaciones. Un individuo humano del Pleistoceno estaba prácticamente condenado a muerte si era expulsado de su clan. Asimismo, los grupos en los que sus individuos no cooperasen intensamente estaban igualmente amenazados de extinción, tanto por el rigor del medio como por la vulnerabilidad frente a otros grupos más poderosos. Lo mismo ocurre con los humanos

actuales separados de toda interacción social. Las abejas y las hormigas, insectos hipersociales por antonomasia, no pueden sobrevivir fuera de sus estructuras de cooperación en las colmenas y hormigueros. De este modo, la interdependencia es un acicate para cooperar, tanto en el caso de individuos egoístas como en el de altruistas.

El problema surge cuando hay individuos que se aprovechan de las ventajas de la agrupación pero no cooperan con los demás miembros del grupo. Conforme se va incrementando aleatoriamente el tamaño de la población crece de forma paralela la diversidad y heterogeneidad de preferencias y rasgos conductuales, y esto reduce la interdependencia de objetivos y disposiciones a la acción común. Dado que la heterogeneidad disminuye la probabilidad de una eficiente acción común, en ambientes de heterogeneidad es más probable que se den tasas de cooperación reducidas, estrategias dominantes y por lo tanto resultados grupales inferiores a los de una población homogénea. En ambientes en los que se mezclan cooperadores y defectores, estos últimos están produciendo una externalidad negativa que perjudica a los demás que sí cooperan. Si los cooperadores no son capaces de reconocer a los defectores, el acicate de la interdependencia se ve disminuido y el sistema como tal pierde fortaleza. El problema no surgiría si todos los individuos fueran cooperadores, pero la conducta de *free-rider* aporta indiscutibles ventajas individuales cuando no se es un cooperador y la explotación de los cooperadores permite una supervivencia parasitaria. Consecuentemente, en poblaciones heterogéneas se dan distribuciones con mayor peso de conductas defectoras que en otras relativamente más homogéneas.

La existencia de preferencias exógenas causadas por herencia genética o cultural más proclives a la cooperación, o por el contrario a aprovecharse de los demás, ha sido estudiada teórica y experimentalmente (véanse Boyd y Richerson 1985; Durham 1991; Ben-Ner y Putterman 1998). Existen individuos que tienden a cooperar (aun en primera jugada de interacciones únicas, y en última jugada de interacciones repetidas) mientras que otros tienden sistemáticamente a comportarse como defectores, e incluso como “castigadores perversos” (Ones y Putterman 2007: 499). En algunos textos se especifican rasgos conductuales propios de una personalidad altruista frente a otras tipologías de carácter egoísta (Frank 1988: 234-236), y asimismo existe un patrón

observado de conducta consistente en el castigo antisocial de los no-cooperadores a los cooperadores que practican la reciprocidad fuerte (Rand y Nowak 2011). Esto no tiene por qué ser un problema si resulta factible detectar a los oportunistas y no-cooperadores en general, penalizarlos y/o excluirlos de futuras interacciones, e incluso construir un sistema institucional de “cinturones protectores” que preserve a los cooperadores de la invasión de oportunistas²¹⁵.

Pero si bien estos supuestos, que constituyen la teoría del enclave cooperativo, pudieron explicar la evolución de la cooperación en las bandas nómadas de cazadores-recolectores del Pleistoceno (y en las bandas endogámicas de primates no humanos de la actualidad), no sirven para explicar la cooperación en las grandes ciudades agrarias neolíticas, y mucho menos en las heterogéneas e impersonales sociedades industriales o post-industriales de la sociedad globalizada. Ello es así porque es difícil de explicar con los argumentos expuestos la consolidación del enclave cooperativo fuertemente cohesionado y “amurallado” cuando las interacciones bilaterales no se repiten o se repiten sólo ocasionalmente (es decir, aparecen continuamente individuos desconocidos para interactuar), y tampoco cuando resulta muy costoso o incluso imposible controlar la incursión de extraños, e/o identificar a los defectores²¹⁶.

Estas dos últimas complicaciones suelen ser frecuentes en las organizaciones de gran dimensión y con una elevada rotación/renovación de sus integrantes. Asimismo, ocurre en organizaciones (grandes o pequeñas, aunque más fácilmente en las grandes debido a su menor densidad relacional) en las que la tecnología, la inexistencia de estándares objetivos, o la dificultad de supervisión y control, limitan las posibilidades

²¹⁵ El propio Axelrod admite que un solo cooperador no podrá invadir nunca a una población establecida de oportunistas. Sin embargo si los cooperadores de *tft* cuentan con algún tipo de organización grupal (que sean endogámicos y no se relacionen demasiado con *outsiders*), pueden no sólo evitar invasiones de oportunistas, sino incluso invadir a una población de defectores (Axelrod 1996: 139). Aunque el autor no lo especifica, siendo así puede hablarse de interacciones discriminatorias, en las que resultan necesarias capacidades de señalización e identificación. Tales son los casos de la segregación espacial, si lleva implícita una identidad social-territorial, o la señalización a través de lo que Axelrod denomina “etiquetas” de reconocimiento (véase pp. 155 y ss., y p. 150). En estos casos «la maquinaria correspondiente a la evolución de la cooperación tiene un retén de enclavamiento» (*ibid.*, p. 168), o sea, aparece un efecto trinquete que en la práctica equivale a un punto histórico de no-retorno.

²¹⁶ Bowles y Gintis (2003: 432) mencionan que los ambientes de grupos con apertura a nuevos ingresos y salidas de individuos limitan notablemente las posibilidades de infligir reciprocidad negativa a los oportunistas, especialmente si éstos tienen facilidad para ser acogidos en otros grupos y no existe comunicación ni posibilidad de reciprocidad indirecta en las relaciones inter-grupales.

de vigilancia e identificación de los miembros del grupo que no cooperan. Si esto es así, cuando se trata de un proceso productivo que requiere acción de equipo, estos aspectos señalados repercuten en la aparición de externalidades que deterioran el nivel de eficiencia grupal, y de ello las posibilidades de existencia y continuidad del equipo.

3.4.2.- El problema de la “visibilidad”

La segunda problemática tiene que ver con las posibilidades fácticas de ejecutar reciprocidad negativa a los defectores, incluso por terceros no directamente perjudicados (reciprocidad indirecta). Dado que la capacidad de observar y el sentirse observados contribuye al refuerzo de la cooperación (Hoffmann, *et. al.* 1998: 344), entonces la cuestión clave aquí es si se puede disponer o no de información accesible y barata, y de buena calidad (nulos o irrelevantes costes de transacción), para que, *cæteris paribus*, los cooperadores puedan protegerse de las externalidades negativas de los defectores:

La idea básica consiste en que un defector no le sea posible escabullirse sin que los individuos a quienes ha engañado puedan desquitarse. La respuesta requiere que el individuo defector no pueda perderse en un mar de anonimato (...) en condiciones de interacción aleatoria y relaciones transitorias en las que el reconocimiento resulta imposible es mucho más probable la explotación. (Axelrod y Hamilton 1981: 1.395).

Se trata pues de un asunto de “visibilidad”, sobre el que conviene subrayar diversos asuntos relevantes:

- A) Por un lado, tal y como exponen Bowles y Gintis (2003: 432) el pasar de un conglomerado de interacciones típico del pequeño grupo idiosincrásico a otro propio de grandes sociedades impersonales conlleva problemas de viabilidad de la cooperación. Ello es así especialmente porque el paso de patrones de cooperación diádicos estables a otros en los que es infrecuente esa estabilidad provoca un incremento de las posibilidades de defección conforme aumenta la expectativa de interactuar con individuos desconocidos. Este hecho deteriora la reciprocidad, y lo hace más conforme mayor es la frecuencia de posibles interacciones con desconocidos. Se trata por tanto de un problema convencional de costes de transacción, ya que:

- a) Cuanto mayor es la cuantía de relaciones potenciales y la incertidumbre asociada a ellas, así como mayor la cantidad de defecciones soportadas por un cooperador, más elevado resulta el coste individual de represaliar a los defectores;
- b) en ausencia de señalización eficaz el coste de encontrar a los cooperadores en la gran amalgama social de individuos heterogéneos se incrementa conforme aumenta el tamaño de la población y la consiguiente incertidumbre. Por otro lado, la solución de búsqueda por tanteo aleatorio no deja de ser arriesgada;
- c) asimismo, el coste de implementar mecanismos conjuntos de coordinación e incentivos, que garanticen el castigo a los defectores, se incrementa también con el tamaño de la población (como también la complejidad y la organización del gobierno de esos mecanismos), y además aparecen problemas adicionales de agencia, financiación y sostenimiento de los bienes públicos;
- d) aunque en las grandes poblaciones heterogéneas se mantiene una preponderancia de interacciones diádicas sin embargo, conforme se incrementa el tamaño poblacional, aparecen situaciones más complejas que requieren una cooperación no diádica, sino multipersonal, típica de la producción conjunta de bienes públicos (como por ejemplo la defensa de la población frente a amenazas externas, reducción de riesgos mediante estructuras de aseguramiento, sostenimiento de las normas sociales beneficiosas para el colectivo, o la difusión de información de buena calidad). De este modo, cuando las poblaciones son muy cuantiosas y heterogéneas, el teorema central de la evolución de la cooperación que plantea Axelrod no ofrece garantías que permitan prever una cooperación abundante y estable (Bowles y Gintis 2003: 432).

Así pues, un ecosistema como la gran sociedad impersonal es un ambiente relativamente “caro” para la evolución de la cooperación, y relativamente barato para la evolución de estrategias no cooperativas.

B) En segundo lugar, es preciso resaltar que la propia estrategia de reciprocidad consiste en la defección, entre otras represalias posibles, por parte de los cooperadores que han sido víctimas del oportunismo. Sin el castigo en la jugada siguiente la cooperación de *tft* no sobreviviría. Pero tienen razón Richerson y Boyd (1998: 76) cuando señalan que en el caso de grandes grupos, en los que no se producen conglomerados estables de interacción entre individuos, esta pauta de comportamiento resulta ineficiente si no es posible diferenciar o separar a los cooperadores de los no cooperadores. Ello es así porque los cooperadores traicionados que no sean capaces de reconocer la identidad de quienes no han cooperado (con ellos o con el grupo), o que no sean capaces de represaliar a los defectores debido a que no repiten jugada con ellos, tenderán a aplicar la represalia de defección de forma indiscriminada en las jugadas siguientes, perjudicando tanto a cooperadores como a no cooperadores²¹⁷. Fowler y Christakis (2010) muestran que el comportamiento, cooperativo o no, se adapta a las experiencias precedentes, aunque las interacciones futuras sean con individuos desconocidos. Andreoni (1988) se ha manifestado en una línea parecida, al concluir que la decadencia paulatina de la cooperación en juegos

²¹⁷ La estrategia Pavlov puede ser confundida con *tft* dado que pasa de la cooperación a la defección si se ha sido traicionado en la jugada anterior. Sin embargo, se trata de una estrategia diferente que puede ayudar a explicar el comportamiento de reciprocidad en contextos de sociedades extensas, o en individuos incapaces de reconocer y clasificar a otros jugadores. Nowak y Sigmund (1993a) la han denominado “comportarse igual si se gana, y cambiar si se pierde” (*win-stay, lose-shift*), y consiste en repetir el comportamiento de la jugada anterior si se ha conseguido un resultado satisfactorio, y cambiar de conducta si ha ocurrido lo contrario. Pavlov es una buena estrategia para reaccionar rápidamente contra la explotación, dado que la defección del otro jugador es interpretada como un castigo, y por lo tanto el estratega pavloviano que comienza cooperando y no es correspondido deja inmediatamente de cooperar. Éste sólo volverá a hacerlo si se topa con cooperadores que demuestran con los hechos que lo son. Advuértase que esta es una salida posible de nuevo a la cooperación cuando existe una elevada incertidumbre. El cooperador que castiga indiscriminadamente pierde el beneficio de la cooperación con otros cooperadores al “enrocarse” en una estrategia de desconfianza y represalia demasiado rígida. Lo mismo ocurre cuando se niega a volver a jugar con alguien que le traicionó en el pasado (conducta de resentimiento). Esa traición pudo ser un error que no necesariamente ha de volver a producirse. El psicólogo Michael McCullough (2008) encuentra en las emociones relacionadas con la conducta de perdón un rasgo instintivo que aporta ventajas adaptativas al individuo que lo porta. Esta actitud de benevolencia y perdón puede haber evolucionado en paralelo al “ojo por ojo y diente por diente” (Fudenberg, *et. al.* 2012; Dreber, *et. al.* 2008).

anónimos repetidos de provisión de bienes públicos, en los que no es posible una represalia individualizada por la vía de la exclusión del disfrute a quien no coopere, se debe a que los cooperadores tienden a represaliar a los *free-riders* del único modo que les es posible: dejando de cooperar. Estos procesos generan comportamientos contagiosos y realimentan tendencias colectivas a favor y en contra de la cooperación (Fowler, *et. al.* 2010). El origen de ello radica en que la cooperación entre egoístas necesita ser condicional, pues de no ser así no aportaría ventaja alguna al individuo. La economía experimental demuestra que sólo una fracción reducida de individuos se comporta defectando sistemáticamente y no coopera nunca en los juegos, por ejemplo, de provisión y sostenimiento de bienes públicos (Ledyard 1995). La mayoría de los individuos comienza cooperando en mayor o menor medida, pero conforme se van repitiendo las jugadas su nivel de cooperación tiende a converger hacia la media. Se ha comprobado además que el nivel medio de cooperación tiende a decaer cuando la jugada se repite en numerosas ocasiones y se aproxima el final del juego. Una posible explicación de estos procesos convergentes puede ser que la condicionalidad de la cooperación se desarrolla a través del aprendizaje social²¹⁸. Los cooperadores condicionales adaptan su estándar de cooperación al comportamiento observado en los otros, aunque partan de preferencias iniciales dispares de contribución cooperativa:

Los cooperadores condicionales cooperan si observan que una cantidad suficiente de los demás también lo hace (...) La razón es que los cooperadores condicionales, cuando se dan cuenta de que entre ellos hay algunos *free-riders* que se aprovechan, tienden a reducir su contribución para no actuar como incautos. (Gächter y Thöni 2005: 303).

Se ha demostrado que la aportación de los cooperadores condicionales obedece a un proceso mental de cognición social que incluye la formación de expectativas en las que se aprecia una clara adaptación imitativa del entorno conductual

²¹⁸ En psicología social está aceptado que la expectativa acerca del comportamiento de los demás influye destacadamente en el comportamiento propio, muchas veces en orden a imitarlo (Darley y Fazio 1980). La experimentación llevada a cabo con juegos de provisión de bienes públicos corrobora igualmente esta hipótesis (Fowler, *et. al.* 2010). El comportamiento en la interacción social depende, entre otras variables, de las creencias de primer orden acerca de las creencias, preferencias, motivaciones, intenciones y comportamientos finales de los demás, lo cual determina un proceso de ajuste gradual de tipo convergente (Snyder y Stukas 1999; Levine 1998).

(Weber y Murnighan 2008). En este sentido, Arrow (1974: 73) ha denominado “expectativas convergentes” al patrón de formación de expectativas y toma de decisiones que tiende a adoptar la estrategia de conducta que espera encontrar en el universo de posibles interactores. De este modo, el comportamiento dominante que acaba imponiéndose es el que se cree más probable encontrar en otros, y si esto es así en toda la población, las expectativas individuales tienden a converger, expandiendo o contrayendo la cooperación y la defección (véase figura 3.8.)

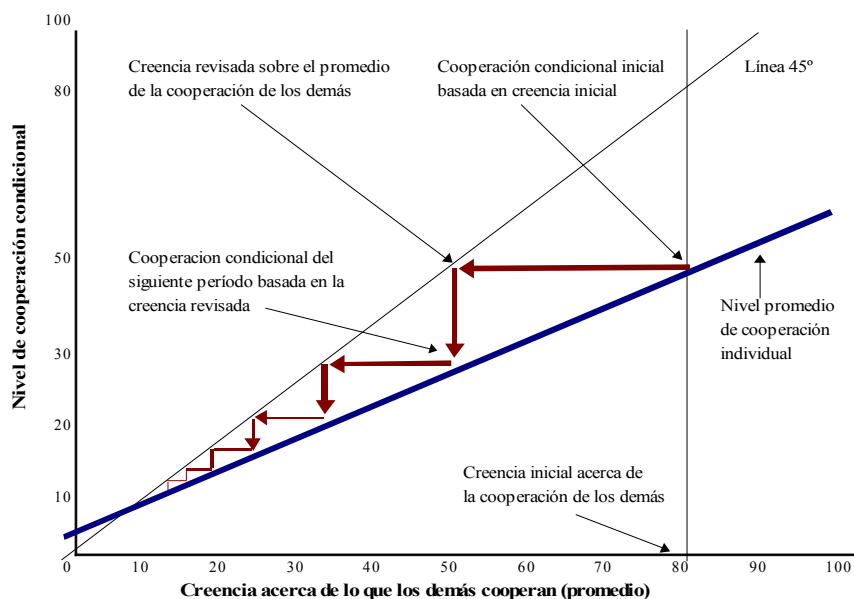
Una conclusión de lo explicado es que un sistema social de grandes números, que se muestre incapaz de generar innovaciones institucionales en orden a reforzar la reciprocidad, tenderá a estabilizarse en niveles mínimos de cooperación porque los cooperadores recíprocos adaptarán su estándar a los niveles inferiores de su función de utilidad. Esto no es incompatible con el resultado de un ecosistema de cooperadores obligados por las circunstancias a comportarse como defectores condicionales. Cuando no es posible represaliar individualmente al defector (p.ej. excluyéndole del consumo o disfrute de bienes públicos, lo cual es harto difícil por la propia naturaleza de este tipo de bienes), este patrón de comportamiento tiende a extenderse, y como afirma Kanazawa (2001: 1.146): «Si todo el mundo actúa como un *free-rider*, el sistema no puede sostenerse y acaba colapsando». Ese colapso es ineficiente en el sentido de Pareto, y al fin y al cabo acaba por perjudicar a todos, como ocurre en los desenlaces de la tragedia de los comunes y del dilema del prisionero.

Howard Margolis (1982) explica las preferencias sociales e interdependientes en un modelo en el que la función de utilidad está compuesta de preferencias individuales así como de preferencias sociales. El patrón de motivaciones del agente estaría integrado por motivaciones egoístas, concernientes al bienestar propio, y motivaciones altruistas, concernientes al bienestar comunitario, bien común, bien público o términos similares. Siguiendo este planteamiento podríamos esbozar una función de utilidad tal que:

$$U_i = U(r_i) + U(rs_i) ,$$

donde r_i es la renta individual o privada del individuo i , y rs_i es su contribución a la renta social (una medida de su aportación a los recursos comunitarios o a un determinado bien público). La renta total del individuo es $r_{Ti} = r_i + rs_i$. Admitimos que los cooperadores se diferencian de los no cooperadores en que para los primeros $U(rs) > 0$, mientras que para los segundos $U(rs) = 0$. Además el agente cooperador contribuye con su aportación a la renta social de manera condicional, tal y como se describe en el gráfico de la figura 3.8., es decir, rs_m tiende a igualarse a $(\sum_{(i:1,N)} rs_j)/(N)$, siendo N la población total. Los cooperadores condicionales prefieren incrementar su renta individual reduciendo sus aportaciones a la renta social cuando se dan cuenta de que su aportación supera la aportación promedio. Esto último implica que $\partial U_i/\partial r_i > \partial U_i/\partial rs_i$ cuando el valor de rs_i en la jugada k supera el promedio. Pero puede admitirse que los cooperadores cuentan con un valor mínimo de contribución rs_{min} por debajo del cual $\partial U_i/\partial r_i < \partial U_i/\partial rs_{min}$. Aportaciones inferiores a rs_{min} no resultan aceptables al individuo en términos utilitarios. Este punto de contribución mínima puede ser denominado *punto de altruismo basal*, o *de altruismo pasivo*. Si suponemos que existen M cooperadores en la población, y que todos ellos tienen un segmento de altruismo pasivo en su función de utilidad, entonces el punto donde la línea azul del gráfico de la figura 3.8. corta la bisectriz será tal que la rs_{min} de los M cooperadores que forman el grupo es la que hace que $rs_{min} = (\sum_{(m:1,M)} rs_m)/(M)$.

Figura 3.8.- Formación de expectativas y deterioro de la cooperación



El ejemplo representado en el gráfico superior (inspirado en Fehr y Fischbacher 2003: 787) ilustra el proceso de formación de expectativas convergentes en un grupo heterogéneo de individuos que realizan aportaciones cooperativas para la financiación de un bien público. En el grupo existen cooperadores y *free-riders* mezclados, y no es posible detectar y discriminar personalmente a los que se aprovechan de los demás, aunque sí se puede sospechar su presencia al darse cuenta que la aportación media resultante en cada tanda sucesiva es inferior a la aportación realizada por el cooperador individual. Como consecuencia de la heterogeneidad aparece la línea de color azul, que representa el nivel promedio de cooperación individual. La línea bisectriz representa promedios de cooperación individual coincidentes con las aportaciones de los cooperadores. En cualquier caso, se supone que los cooperadores llevan a cabo una cooperación condicional, esto es, parten de una expectativa acerca del nivel de cooperación de los demás y deciden cooperar en esa misma medida (80, en el ejemplo), o sea, cooperan más o menos según sea su percepción del nivel de cooperación agregado. Como quiera que el resultado promedio de todas las aportaciones será inferior (la bisectriz tiene más pendiente que la línea de nivel promedio), los cooperadores, que son condicionales, reajustarán su expectativa sobre la aportación de los demás (hasta el nivel de 50), pero aquí se volverá a repetir el proceso, pues constatarán que la aportación media es inferior a 50 (se sitúa en algo más de 30). El proceso conduce a una convergencia a la baja en la que la cooperación llegará a ser testimonial, y esto ocurre cuando se produce la intersección entre la línea bisectriz y la línea que marca el nivel promedio de cooperación. En este punto se da un equilibrio estable. Según Fehr y Fischbacher (2003) la conclusión más interesante de este modelo consiste en que un grupo, aunque esté poblado por una mayoría de cooperadores, se mostrará incapaz de evitar la reducción paulatina de la cooperación, a pesar de que los egoístas aprovechados son una minoría. La razón de que esto sea así se explica por el anonimato de los *free-riders* acompañado del mecanismo de formación condicional de expectativas, única forma de castigo genérico que pueden aplicar los cooperadores ante su incapacidad de discriminar y aplicar reciprocidad negativa específica a los individuos que se aprovechan de los demás.

Un cooperador que haya sufrido una historia relacional de defecciones tenderá a extrapolar esos resultados y a descontar futuras defecciones en encuentros con desconocidos. Hardin (2002: 113) lo explica con las siguientes palabras:

Cuando me encuentro con alguien nuevo con quien pretendo o tengo que tratar, puedo iniciar el contacto con diversas actitudes (...) en gran medida (condicionadas) por extrapolación de encuentros pasados con otras personas (...) mi experiencia moldea mis expectativas sobre la fiabilidad de los demás. (Paréntesis añadido).

De este modo, existe un efecto trinquete que viene explicado por que las expectativas formadas sobre el comportamiento ajeno a la larga se convierten en profecías auto-cumplidas. Consecuentemente, *tft* tenderá a degenerar en sociedades de grandes números, pues la incertidumbre hará proliferar la defección, junto con la desconfianza y la astucia de los oportunistas²¹⁹.

En economía esta hipótesis del deterioro paulatino de la ética de la reciprocidad conforme se incrementa la despersonalización de las interacciones con el tamaño del grupo, ha sido tratada hace tiempo en varios trabajos por Buchanan, y puede sintetizarse en la siguiente hipótesis sobre la inconsistencia aparente entre la disposición cooperativa del individuo humano y el tamaño de la población en la que éste interacciona:

La fuerza de los principios ético-morales en la motivación del comportamiento depende directamente del tamaño de la comunidad dentro de la que tiene lugar la acción. Todo lo demás igual, cuanto menor es el número de personas con las que un individuo interacciona, mayor es la probabilidad de que tienda a comportarse de acuerdo con algo parecido al imperativo kantiano: en nuestra terminología, esto último supone que el individuo contribuirá a la comunidad con más bien público. (Buchanan 1978: 365)²²⁰.

²¹⁹ De hecho, la literatura de economía experimental sobre contribución a bienes públicos coincide en que la cooperación entre individuos seleccionados aleatoriamente tiende a reducirse de forma drástica con el transcurso de las iteraciones del juego (Ledyard 1995). En experimentos de provisión de bienes públicos llevados a cabo con grupos homogéneos y heterogéneos se observa que los grupos formados exclusiva o mayoritariamente con cooperadores mantienen elevados resultados de cooperación en sucesivas iteraciones hasta que llegan las últimas jugadas, mientras que los grupos formados sólo con defectores, o mayoritariamente con ellos (mezclados con una minoría de cooperadores) dan resultados claramente degenerativos una vez transcurridas las primeras jugadas (véase Gächter y Thöni 2004; y Ones y Putterman 2007).

²²⁰ Esta tendencia degenerativa de la cooperación y de la ética de la reciprocidad en las grandes sociedades modernas ha sido planteada también por la etología humana. Véase Eibl-Eibesfeldt (1996).

C) La evolución de la cooperación reforzada por lo que podemos llamar una ética de la buena reputación requiere transparencia en los comportamientos, es decir, una información accesible y de buena calidad, y suficientemente difundida, para que los comportamientos vengan a ser información de dominio público. De no ser así los cooperadores no contarían con el apoyo y la confianza de otros cooperadores, y los defectores no podrían ser detectados y castigados, o simplemente evitados (Huck y Tyran 2007; Zucker 1986). Además, la búsqueda “hipócrita” de una buena reputación por parte de individuos, que utilizan la etiqueta de la reputación como medio fraudulento para obtener ventajas, se verá facilitada en contextos de anonimato o de falta de publicidad, o credibilidad, de este marcador diferencial²²¹. Así, en los estudios de economía experimental llevados a cabo por Hoffman y sus colaboradores con el juego del dictador, se ha contrastado que los mismo individuos se muestran más o menos cooperativos en función del grado de publicidad de sus acciones; por eso, los autores concluyen:

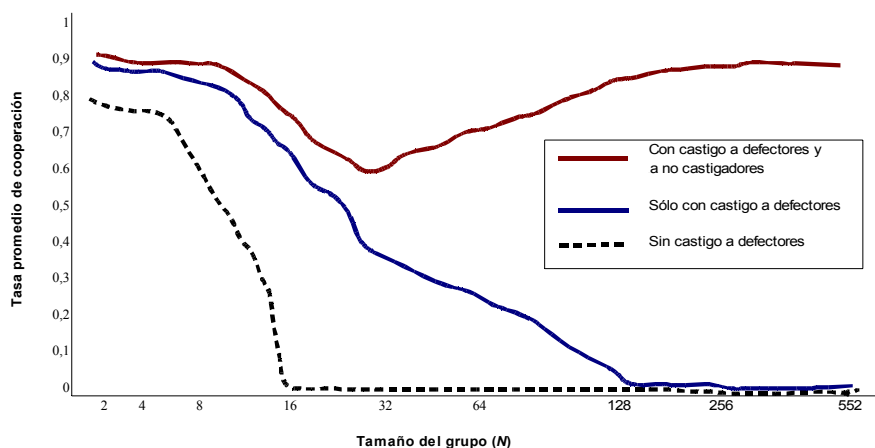
... la gente actúa como si le importaran los demás porque su situación mejora con la reputación resultante. Sólo en condiciones de aislamiento social tienen poca fuerza las preocupaciones por la reputación. (Hoffman, *et. al.*, 1994: 659).

La fuerza de las normas grupales y del control social en el incentivo al autocontrol del comportamiento individual es un hecho admitido por la psicología social y ha sido demostrado experimentalmente (Mittone y Ploner 2011; Sausgruber 2009). Pero para que la reputación actúe como incentivo en individuos oportunistas requiere que éstos se sientan observados y controlados (Loewenstein 1999a: F31-F32). Adicionalmente, los ambientes en los que además de existir publicidad sobre el carácter e historial cooperativo de los agentes existen cooperadores no sólo capaces de evitar a los oportunistas, sino además de “ir a buscarlos” e infligirles un castigo lo suficientemente incentivador como para “hacerles pagar” por su externalidad negativa (*strong reciprocation*), generarán un refuerzo de la cooperación vía incentivos automáticos, y ese refuerzo tiene efectos sistémicos, como se ha demostrado en la literatura experimental (véase figura 3.9.)²²²

²²¹ Véase el teorema del “niño malo” de Becker (1987) para un análisis sobre el comportamiento optimizador de un aparente altruista, que en realidad actúa hipócritamente en un contexto de información asimétrica.

²²² Trabajos experimentales sobre el comportamiento de castigo al defector o *free-rider*, y su incidencia

Figura 3.9.- Relación entre nivel de cooperación y tamaño poblacional en experimentos con dilemas sociales



Fuente: Fehr y Fischbacher (2003: 790).

El gráfico superior contiene los resultados del experimento llevado a cabo por Fehr y Fischbacher (2003) basado en la metodología de Boyd, *et. al.* (2003) para el juego del dilema del prisionero multipersonales con conflictos intra-grupo y diferentes grados de castigo altruista. No obstante, los resultados son semejantes en el caso de otros dilemas sociales multipersonales, como el juego de provisión y sostenimiento de bienes públicos con una función lineal de pagos individuales del siguiente tipo:

$$y_{i,N} = (P - C_i) + a \cdot \sum_{(j: 1,N)} C_j ;$$

donde y es el pago individual que el jugador i obtiene del juego; j es el miembro componente de la población que incluye a i ; N es el tamaño poblacional; P la cantidad de unidades monetarias con las que cuentan todos los jugadores antes de jugar; C la contribución individual del individuo al bien público (que no incluye el valor del castigo que pueda infligir si se trata de un *strong reciprocator*); y a el porcentaje de bien público del que se beneficia el individuo. En un juego de estas características se puede definir la tasa promedio de cooperación como $T = (\sum_{(j: 1,N)} C_j) / (N)$. El experimento puede ser realizado con tres supuestos distintos: (1) sin castigo a los *free-riders* y/o defectores (línea negra de trazo discontinuo); (2) sólo con castigo a los *free-riders* y/o defectores (línea de color azul); y (3) con castigo tanto a los *free-riders* y/o defectores, como a los cooperadores que no castiguen a los *free-riders* y/o defectores (línea de color rojo). Como se puede visualizar, la imposibilidad de represaliar a los individuos que no cooperan un mínimo exigible (según una norma grupal, por ejemplo) se traduce en unos resultados pobres en cooperación conforme aumenta N y con ello el número de oportunistas errantes: en este caso, en grupos con una población superior a 16 individuos, la tasa promedio de cooperación (T) cae drásticamente a cifras cercanas a cero conforme aumenta la población. Los resultados son bastante mejores cuando existen castigadores altruistas que aplican un castigo suficientemente costoso al oportunista; en este caso T cae con menor rapidez y tiende a cero para tamaños grupales superiores a 128 individuos. Los mejores resultados ocurren cuando no sólo existen castigadores altruistas, sino que además se puede castigar a cooperadores que no castiguen a los oportunistas. En este caso T toca su mínimo con N en torno a 30, para después recuperarse paulatinamente hasta cifras de T situadas entre el 70% y el 80% con un tamaño poblacional de más de 500 individuos. Estos modelos se han utilizado para demostrar la viabilidad de la selección cultural grupal, en los que las normas de

positiva en el nivel de cooperación grupal, son los de Ostrom, *et. al.* (1992), Fehr y Gächter (2000a y 200b), Hoffman, *et. al.* (1998), Gächter y Thöni (2005), Anderson y Putterman (2006), y Ones y Putterman (2007). Sobre la conducta de represalia ante ofertas “injustas” en el juego del ultimátum puede verse Davis y Holt (1993). Sobre la represalia en el juego del ultimátum y el juego del dictador puede consultarse Camerer y Thaler (1995).

cooperación reforzadas por instituciones que castigan suficientemente a los no-cooperadores posibilitan la evolución de la cooperación no vía exclusiva de premio a la buena reputación, sino, y especialmente mediante, castigo a la mala reputación asociada a no cooperar suficientemente. Así, los cooperadores y los vengadores altruistas tendrán garantizada su supervivencia y replicación si representan una proporción elevada respecto al total de la población. De ahí el coste individual de aplicar castigos cae de forma significativa. Por su parte, la replicación de las conductas cooperativas no necesariamente ha de ser por la vía genética (neodarwiniana). Es suficiente con que los individuos estén capacitados para el aprendizaje social a través del adoctrinamiento conductista o la imitación de las conductas de éxito. Estos modelos de evolución cultural permiten explicar procesos de cooperación a gran escala y avances rápidos en el tiempo del patrón conductual cooperativo. Un modelo de este tipo permite conjeturar la cooperación fuera del apiñamiento axelrodiano en poblaciones de gran tamaño que superan el número de Dunbar.

Sin embargo, la figura del *strong reciprocator* (una especie de superhéroe moderno) es la de un cooperador altruista dado que, en un modelo que no tiene en cuenta el factor replicación, asume personalmente costes que sólo le serían compensados si ello le permitiera ganar una sólida reputación pública recompensada por gratitud, selección favorable de otros cooperadores, o sistemas institucionalizados de prestigio y reconocimiento social²²³. De no ser así, su actuación altruista no le reporta ninguna ventaja adaptativa, y por ello la estrategia de vengador justiciero tendrá grandes dificultades para sobrevivir a largo plazo. Admítase que el vengador justiciero puede toparse con oportunistas capaces de responder contraatacando agresivamente a su castigo, y ocasionándole pérdidas que él no podría externalizar (Rand y Nowak 2001). La solución sería recompensar colectivamente a esta figura, dado que está desempeñando *de facto* un servicio público de interés general (véase Bowles, *et al.* 1997: 7-10). La figura del policía puede de este modo emerger en este tipo de ambientes, pero en una sociedad sin Estado, muy poblada y con carencia de información, volvemos a encontrarnos con los problemas de provisión y sostenimiento de bienes públicos señalados por Olson (1971[1965]), e incluso

²²³ Bowles, *et al.* (1997) utilizan la etiqueta de *homo reciprocans* para aludir a este patrón de comportamiento en contraposición al *homo oeconomicus* de la economía neoclásica. El *homo oeconomicus* no actúa en pos de la justicia (*reciprocal fairness*) aplicando venganza por su cuenta y riesgo, e incluso contra agravios ocurridos a terceros, salvo que obtenga un beneficio por ello. Sin embargo, la figura del *homo reciprocans* es la de un vengador que aplica castigos morales, que asume costes personales sin expectativa de un beneficio (de supervivencia) posterior, incluso para represaliar agravios realizados a terceros (*strong reciprocity*). Se puede argumentar que tal comportamiento es una cuestión de preferencias y maximización de utilidad, pero lo relevante aquí es cómo pueden existir este tipo de preferencias “irracionales” en ambientes de escasez y competencia, es decir, ¿cómo pueden haber evolucionado preferencias que supuestamente conllevan un coste de oportunidad muy alto en términos de eficacia?

las sociedades con Estado tampoco son ajenas a esta problemática (Olson 1986; Friedman y Friedman 1984; Buchanan 1965b)²²⁴.

3.5.- Conclusiones

La respuesta a la pregunta planteada en la introducción de este capítulo, sobre si resultan suficientes las conductas de reciprocidad y egoísmo prudencial para que la cooperación evolucione en sociedades complejas y no estacionarias, es negativa. Tanto la administración de incentivos a través de una autoridad central dotada de poder, al estilo de la teoría de la organización neohobbesiana, como las diversas formas de reciprocidad analizadas (*tft* y *strong reciprocity*, directa e indirecta) en orden a inhibir la conducta no cooperativa en contextos de poder descentralizado, no garantizan la evolución de la cooperación en ambientes complejos, dinámicos e inciertos como los que caracterizan, y han caracterizado históricamente en mayor o menor grado, una proporción significativa de la interacción social humana.

La teoría del refuerzo de la cooperación vía incentivos y capacidad sancionadora de una tercera parte no garantiza que esa tercera parte cumpla su función, dado que exige un supuesto motivacional, el altruismo, que no es posible asegurar en un individuo que, al fin y al cabo, es representativo de una población que se supone no cumple dicho supuesto.

Prescindiendo pues de la teoría del Leviatán filantrópico hemos analizado las posibilidades de evolución de la cooperación en poblaciones con poder descentralizado. La conclusión es que la cooperación condicional entre individuos basada en la reciprocidad (directa e indirecta) es una EEE siempre y cuando los cooperadores se apiñen en grupos cuya estructura, dimensión e instituciones estén adaptadas a las capacidades cognitivas de la especie. Dichas capacidades aluden principalmente a las facultades memorísticas, de reconocimiento de cooperadores con los que se cooperó con éxito en el pasado, así como defectores a los que se ha de aplicar represalia o eludir interactuar con ellos. Ello supone que el apiñamiento ha de ser limitado en número de

²²⁴ Yamagishi (1986) atribuye a esta posibilidad la consecuencia del problema central de la acción colectiva, que es el fenómeno *free rider*, pero en una categoría de segundo orden; es decir, el control y castigo de los explotadores del grupo deviene en sí mismo un problema típico de acción colectiva.

individuos (es decir, cognitivamente “finito”), estable y duradero, lo que garantizaría una interacción social densa en términos de complementariedad cooperativa (homogeneidad conductual, en pos del logro de resultados interdependientes) y visibilidad (buena información sobre la reputación de los demás individuos). Con estas premisas se ha demostrado que se cuenta con las bases necesarias para que la cooperación, al ser EEE, pueda evolucionar espontáneamente enfrentada a otras estrategias (Nowak, *et. al.* 2004).

La literatura especializada suele caracterizar estos enclaves, o agrupaciones aptas para la cooperación social, como poblaciones de pequeña dimensión, al estilo de las familias y bandas de cazadores-recolectores del Pleistoceno, y por extensión las comunidades “amuralladas” de las sociedades tradicionales. No obstante, el tamaño poblacional óptimo es una variable dependiente de las capacidades cognitivas de los cooperadores. En el caso humano estos enclaves de cooperación tienen un tamaño de equilibrio que podemos considerar muy reducido si lo comparamos con las estructuras de cooperación empíricas propias de la especie. Entre estas últimas se incluyen las grandes ciudades del pasado, y la compleja y ampliamente poblada sociedad global de la actualidad. Sin embargo, el tamaño de los enclaves eficientes, cifrado entre 150 y 600 individuos, es el que han tenido los grupos humanos de referencia durante la mayor parte de la historia evolutiva de la especie. Cuando los grupos superan esta dimensión poblacional -admitiendo que en ellos habitan cooperadores y no cooperadores mezclados y sin identificar, y cierto nivel de opacidad informativa- empiezan a manifestarse tendencias degenerativas en la cooperación. En los modelos que han sido analizados dichas tendencias tienen que ver con que los cooperadores, para protegerse de las defecciones, pueden llegar a reducir sus niveles de cooperación llegando incluso a castigar con defección indiscriminada, o a dejar de interactuar con desconocidos. En circunstancias en que se produce incremento dinámico de la población, con individuos heterogéneos, móviles e interacciones aleatorias, y si no existen *strong reciprocators* (altruistas pro-activos), o están insuficientemente representados en la población, aparece un umbral crítico a partir del cual la cooperación degenera. La razón es que al pasar del apiñamiento viscoso a una sociedad de desconocidos las interacciones se transforman en juegos de única jugada, y esto nos remite a juegos no cooperativos como los planteados

por la teoría de juegos tradicional (Neumann y Morgenstern 1953[1944]), y a los modelos microeconómicos de no-cooperación inspirados en la tradición neohobbesiana (Olson 1971[1965]).

La opacidad en sociedades con una cantidad de población lo suficientemente numerosa y heterogénea constituye un destacado problema tanto para el refuerzo de la cooperación vía reputación como para el refuerzo por la vía de una policía pública. Los ambientes de anonimato tienden a degenerar hacia la impunidad, lo cual juega dinámicamente en detrimento de la cooperación y a favor del oportunismo. *Cæteris paribus*, conforme crece la población y el número potencial de jugadores con los que un individuo puede interactuar, se requerirán entramados institucionales eficaces para sacar a la luz la información que permita hacer de la reputación una estrategia cooperativa viable. En el caso contrario, ni el mecanismo conductual de la reciprocidad directa, ni el de la indirecta (incluida la venganza costosa aplicada sobre terceros), permiten explicar órdenes sociales cooperativos evolutivamente estables. Tal y como han señalado Bowles, *et. al.* (1997: 6), y en sintonía con otros experimentos llevados a cabo (Ostrom, *et. al.* 1992; y Fehr y Gächter 2000a y 2000b), las características poblacionales y estructurales de los órdenes cooperativos estables consisten en un tamaño relativamente pequeño -en comparación con las sociedades actuales- y una estructura social impermeable a la entrada y la salida de individuos, que garantice la estabilidad poblacional. En estos entramados se genera una alta densidad relacional (todos los individuos pueden supervisar sin costes significativos a todos los demás, ejercer reciprocidad negativa a los defectores y positiva a los cooperadores más generosos), y la red de interrelaciones tiene la particularidad de una frecuente reiteración en las jugadas y entre jugadores. Tales circunstancias de viscosidad social permiten hablar de “familiaridad” sin necesidad de lazos familiares, como en el “grupo” descrito por Konrad Lorenz.

Los modelos con esta singularidad, inspirados en la teoría del altruismo recíproco de Trivers (1978[1971]), resultan explicativos de una cooperación más extensa (con una superior distancia genética) que la explicada por los modelos de selección de parentesco (Hamilton 1964). Sin embargo, dichos modelos no permiten, en

principio, ofrecer una explicación evolutiva a la cooperación entre desconocidos sin previa vinculación y en juegos de jugada única. En ausencia de lealtad y confianza, y ante un encuentro único, la expectativa racional de los cooperadores es la desconfianza, y la de los egoístas racionales es no cooperar, y de ahí el cruce de estrategias dominantes que conduce resultados no-cooperativos.

Tampoco escapan de esta problemática los planteamientos teóricos que se apoyan en la hipótesis de selección de grupo y selección cultural como refuerzos de las conductas de reciprocidad fuerte aplicada a quienes incumplen la regla de reciprocidad. En efecto, como muestran Boyd, *et. al.* (2003) en su modelo de cooperación en grupos de individuos no emparentados genealógicamente, la cooperación sólo obtiene éxito evolutivo si existe algún mecanismo de refuerzo institucional de la figura del *strong reciprocator* (que vendría a ser el rol de un proto-Estado espontáneo), y aún así no queda garantizada la viabilidad de la cooperación en sociedades integradas por más de 600 individuos. Obviamente, las sociedades actuales son mucho más pobladas y, sin embargo, en ellas abunda la cooperación despersonalizada con desconocidos y en interacciones de jugada única. Asimismo, en estas sociedades despersonalizadas de grandes números es relativamente notoria la cooperación en ambientes desprovistos de regulación contractual, salvaguardas jurídicas o intervención coercitiva de un Estado-policía.

Russell Hardin (2002: 108-110) se refiere a las grandes sociedades extensas actuales como «contextos de cooperación en ausencia de confianza», en los que las organizaciones formales, los Estados y el mercado llevan a cabo funciones institucionales de refuerzo de la cooperación, y lo hacen suministrando información y protegiendo a los cooperadores de interacciones con oportunistas (lo cual incluye su persecución y castigo institucionalizados). Es una vuelta, al menos en parte, hacia las premisas conductuales y conclusiones del modelo microeconómico neohobbesiano de poder centralizado, o sea, a una cooperación no espontánea y sólo posible con la presencia de una jerarquía ejecutiva que ejerce monopólicamente el rol de *strong reciprocator*. No obstante, vuelve a venir a colación la advertencia de North a este respecto, extensible a toda estructura de gobierno con poderes institucionalizados:

... si el Estado posee la fuerza coercitiva, entonces quienes rijan el Estado usarán esa fuerza en su propio interés a expensas del resto de la sociedad. (North 1993: 81-83).

En todos estos casos los individuos psicológica y motivacionalmente egoístas juegan la estrategia defección, y por ello, en interacciones de acción colectiva, los equilibrios resultantes ofrecen los peores resultados desde el punto de vista individual y social. Así, la paradoja del altruismo planteada por Darwin no es tal, y se transformaría en una paradoja del egoísmo. En efecto, la paradoja del egoísmo plantea otros problemas, dado que el actor racional de la teoría de juegos y el *homo oeconomicus*, siendo como se les representa racionales, están paradójicamente abocados a conseguir los pobres resultados de un “tonto racional” (*rational fool*) carente de inteligencia social, y no el éxito atribuible a un astuto egoísta que siempre juega con la intención de sacar algún provecho (Sen 1977). El tonto racional no es un individuo humano real, sino una ficción poco representativa de las facultades cognitivas evolucionadas en la especie *Homo sapiens*, entre las que caben destacarse la disposición empática y la teoría de la mente. Así, en interacciones con la estructura del dilema del prisionero, y por extensión en toda acción colectiva, los perspicaces jugadores racionales consiguen peores resultados que si actuaran de forma irracional, algo que ha llevado a Howard Raiffa (1992: 172) a concluir que dos tontos de manual (los irracionales) obtienen mejores resultados que dos “listos” (los racionales). Naturalmente, esto no es sólo una paradoja, sino además un *non sequitur* desde el punto de vista evolutivo ya que, prescindiendo de las nociones doctrinales sobre la racionalidad, la naturaleza no selecciona patrones epistémicos y conductuales que conducen sistemáticamente a callejones sin salida o paradojas del tipo perder-perder. La cooperación, por lo tanto, no puede ser una anomalía, sino una solución adaptativa a la paradoja del egoísmo. De este modo la pregunta sigue en el aire: ¿hay una mano invisible oculta que maneja el comportamiento en pos de la cooperación?

Las sociedades humanas de las grandes civilizaciones post-neolíticas, y no digamos las grandes sociedades comerciales de la actualidad, no cumplen con las condiciones descritas del apiñamiento cooperativo. Esto supone un reto a la teoría de la cooperación dentro del modelo de reciprocidad entre egoístas racionales. En ambientes

determinados a la interacción con extraños la cooperación tendería a ser mínima (altruismo pasivo), y a la larga un comportamiento disfuncional destinado a desaparecer. Sin embargo, no podemos decir que las grandes poblaciones con un tamaño superior al enclave óptimo se caractericen por la sistemática no-cooperación o por una cooperación testimonial, y ello tanto en sociedades arcaicas de Estado mínimo, como en las actuales sociedades de Estado disfuncional y Estado fallido. Es más, un rasgo definitorio del éxito de la gran sociedad comercial es una contra-intuitiva abundancia de cooperación, incluso con extraños, y aun asumiendo que el incentivo a no cooperar es más motivante en muchas ocasiones para el actor racional. ¿Cómo se explica entonces el que se produzca una cooperación relativamente significativa más allá del enclave de interacción densa?

Como se ha explicado en el apartado 3.2.1., una de las capacidades adaptadas en la especie humana es la de cálculo, que unida a la evolución institucional del dinero, permiten articular un entramado de relaciones sociales de cooperación viable incluso en sociedades de grandes números. Asimismo, el altruismo fuerte deja de ser disfuncional si constituye una señal creíble para individuos que están dispuestos a recompensar al altruista por su conducta con oportunidades para que incremente su eficacia (Lotem, *et. al.* 2003; Gintis, *et. al.* 2001). Ello significa igualmente que la señal de equitatividad plasmada en las transacciones con dinero, y la posesión del mismo, constituyen señales que informan sobre la reputación en órdenes sociales monetarios, en los que el dinero tenga un valor estable y su acumulación no se asocie generalmente a las conductas de defección.

En los siguientes capítulos se analiza la cuestión de si es posible explicar la cooperación desde premisas alternativas al arquetipo utilitarista del individualismo motivacional, modelo estándar de actor racional, o egoísmo psicológico. Esto invita a poner a prueba la hipótesis de que los individuos -o al menos una parte de ellos- tienden instintivamente a cooperar con sus afines, e incluso a ser altruistas con ellos, y además se hallan adaptados para defenderse de la amenaza de depredadores y oportunistas. Esto invita a poner en valor las teorías defendidas por Darwin en *El origen del hombre*, la tradición evolucionista inspirada en el pensamiento de Kropotkin, la psicología social de

Adam Smith, y las formulaciones del s. XX sobre la selección grupal y la evolución y coevolución por cooperación simbiótica.

Si los cooperadores existen empíricamente ha de haber una “coreografía” o “mano invisible” ordenadora que regule y armonice los comportamientos individuales. Descubrir en qué consiste esa coreografía aparentemente oculta y explicar de dónde procede son objetivos a abordar en los siguientes capítulos.

Capítulo 4.- La coreografía oculta de la cooperación: causas últimas y sistemas de refuerzo del comportamiento altruista

Mejor dos que uno, pues obtienen mayor recompensa en sus fatigas. Porque, si caen, uno levantará al otro. Pero, ¡ay si uno cae sin tener a nadie que lo levante! Si dos se acuestan juntos, se calientan; pero uno solo, ¿cómo se calentará? Uno puede ser vencido; dos, en cambio, resisten mejor; pues no se rompe fácilmente una cuerda de tres cabos. (*Ecl* 4, 9-12).

Los biólogos se preguntan por qué los organismos hacen esto o aquello. Con frecuencia se preguntan por qué los organismos se reúnen formando sociedades. No se preguntan –y deberían hacerlo– por qué la materia viva se agrupa primero para formar organismos. (Dawkins 1993[1976]: 305).

Los epítetos de *sociable, simpático, humanitario, compasivo, agradecido, amistoso, generoso, benefactor* u otros equivalentes, son conocidos en todas las lenguas y expresan de manera universal el mérito mayor que la *naturaleza humana* es capaz de alcanzar. (Hume 1993[1751]: 39; cursivas en el original).

... cuando ha ocurrido un accidente o alguien está en peligro, no nos disponemos a calcular cómo afectará el acto de ayudar a la promoción de los objetivos personales. Actuamos inmediatamente porque existen ciertas reglas que seguimos. Esas reglas de comportamiento dan confianza a las personas acerca de lo que pueden esperar unas de otras. (Sen 1989: 145).

Si aceptamos que la propensión humana a cooperar en grupos no es una mera construcción social ni el resultado de la elección racional, y que la cooperación posee una base natural o genética, uno se pregunta cómo llegó ahí. (Fukuyama 2000: 217).

La gente no hace amigos, / los reconoce. (Vinicius Morais, “Amigos”).

4.1.- Introducción

La psicología de las últimas décadas ha contribuido a reconocer el comportamiento altruista depurándolo de posibles motivaciones utilitarias. Es el caso, por ejemplo, de la búsqueda de una gratificación psicológica derivada del acto de ayudar a otro, sin ninguna motivación referida a los intereses o necesidades del individuo que recibe la ayuda. Estas investigaciones han demostrado que hay personas que se comportan de acuerdo al canon de un altruismo auténtico, es decir, sin que su motivación sea una ganancia hedónica directa o una expectativa de reciprocidad presente o futura en beneficio de sí mismas o de sus parientes. Además estos individuos actúan soportando un coste o sacrificio no externalizable, y lo hacen incluso en situaciones en que les es posible escaquearse sin recibir ningún tipo de castigo, amonestación, sanción moral o económica, salvo el que pueda ser sentimiento íntimo de la culpa (Batson 1991, y 1992).

Las investigaciones llevadas a cabo por Milgram (1974) sobre la influencia en la conducta de los innatismos del conformismo con el grupo y la obediencia jerárquica demuestran que, aunque estas propensiones pueden llegar a inhibir la aparición de sentimientos morales y emociones prosociales en determinadas circunstancias de tratamiento cruel a otros, sin embargo, la relajación del posible castigo al disidente, que se niega a infligir un tratamiento doloroso por orden de un superior, y sobre todo la contemplación directa de la víctima por aquél, facilitan la activación de una auténtica preocupación empática por la vida y el bienestar de la víctima. Esto ocurre incluso si se trata de individuos completamente desconocidos. La obediencia de la orden de producir un daño a otro se reduce hasta un 70% cuando el individuo interviene directamente en causar el daño; y si la orden de dañar al otro proviene de una figura no autoritaria la obediencia se reduce prácticamente a cero. Cuando los individuos participantes en el experimento se sienten escasamente vigilados por la autoridad o por otros miembros del grupo aumentan y se generalizan las conductas de desobediencia y escaqueo. El interés por la vida y el bienestar de un desconocido es una motivación suficiente para compensar la satisfacción que pueda aportar sentirse miembro de un grupo u obedecer a un jefe. Tal y como manifiestan Bowles y Gintis (2011: 230):

... la naturaleza de las preferencias reveladas en los experimentos de conducta y en otros estudios sobre comportamiento humano es consistente con el punto de vista de que el genuino altruismo -la voluntad de sacrificar los intereses de uno mismo para ayudar a individuos no emparentados- es la causa próxima de gran parte de la cooperación humana.

Una experimento adicional que refuerza la tesis de la existencia de una motivación intrínseca en el genuino altruismo es la autorregulación de la disposición a mentir a través de la aversión a la mentira (propia y ajena). Una cantidad sorprendentemente elevada de participantes en el experimento de López-Pérez y Spiegelman (2013), que asciende al 40%, se comporta honestamente a pesar de que decir la verdad tenga consecuencias pecuniarias negativas o conlleve la exposición a un peligro²²⁵.

El hipotético buen samaritano de la famosa parábola evangélica (*Lc* 10, 29-37) atendió con sus recursos personales al desvalido del camino, y lo hizo sin expectativa de reciprocidad. Por el contrario, el sacerdote y el levita pasaron de soslayo esquivando al desamparado para evitar así la angustia empática y el malestar psíquico activado por la emoción de la culpa²²⁶. El samaritano obró un genuino altruismo, mientras que los otros dos, que evitaron el encuentro anónimo con el necesitado, tal vez sí lo hubieran ayudado de haber sido pública su conducta. Dado que su motivación no se refería al desvalido, como demostraron en la parábola, pudieran haberlo hecho sólo para salvar su reputación ante los demás.

Si el altruismo es una señal de calidad, es preciso que se pueda identificar al verdadero altruista y diferenciarlo del hipócrita. La cuestión, como se explicará, no es en absoluto baladí. La neurociencia ha descubierto que en el cerebro se activan circuitos diferentes cuando se actúa por evitar la culpa o la vergüenza, y cuando se actúa por un genuino interés por el bienestar de una persona necesitada (Lamm, *et. al.* 2007; Singer y Lamm 2009). El segundo caso es un altruismo psicológico, el primero sin embargo es egoísmo psicológico. La identificación psíquica del sujeto con otros sujetos físicos,

²²⁵ Se está estudiando la posibilidad de que exista un dispositivo congénito autorregulador del comportamiento moral en las personas sanas, que no aparece o que funciona patológicamente en los trastornos de personalidad antisocial y otras patologías sociopáticas (Sommer, *et. al.* 2010).

²²⁶ Una reflexión dentro del análisis microeconómico neoclásico sobre las estrategias emocionales de los mendigos para obtener utilidad a costa de los transeúntes, y la distinción entre el auténtico altruismo (motivado en la emoción del amor) y el altruismo hipócrita (motivado en la emoción de la culpa o la búsqueda de reputación) puede leerse en Becker (1996).

imaginarios o abstracciones representativas de conjuntos de individuos, resulta suficientemente motivadora como para producir acciones de altruismo genuino en la especie humana (Tyler y Blader 2000; Richerson y Boyd 2001). Esto quiere decir que existe el altruismo psicológico y que la ciencia ha demostrado su veracidad empírica.

Numerosos estudios concluyen que el altruismo humano es un rasgo epigenético que puede ser reforzado culturalmente si bien, como tal, su existencia no depende de la cultura sino a la inversa. Las decisiones tomadas de manera impulsiva, con una heurística rápida y frugal, tienden a ser más propensas al comportamiento altruista que las mismas decisiones tomadas de una forma lenta, deliberada y reflexiva (Rand, *et. al.* 2012). Esto informa de la existencia de un patrón innato de acción altruista en la especie humana que funciona por defecto. Efectivamente, el que podríamos llamar “instinto de cooperación” está indisolublemente unido a la forma de vida estructurada en torno a la acción colectiva y necesidad de actuar agrupadamente. Tal instinto es de origen filogenético, y se manifiesta ontogenéticamente con carácter previo a otras adaptaciones como son las emociones y desarrollos cognitivos para la cooperación condicional, la detección de cooperadores y defectores, la reciprocidad negativa, la vinculación cultural e incluso la propia selección de parentesco (Trivers 2004). Al contrario, los comportamientos egoístas del tipo medios-justificados-por-un-fin, que repercuten en beneficio propio a costa de los medios o fines de otros, aparecen típicamente, y de manera sistemática, en individuos con disfunciones o discapacidades en el área cortical prefrontal. Estas personas no experimentan sentimientos de culpabilidad cuando incumplen normas morales básicas, y su conducta se considera clínicamente patológica (Koenigs, *et. al.* 2007).

El altruismo como rasgo de la especie humana se aprecia en todas las culturas y en diferentes situaciones empíricas de nivel de renta y desarrollo (Aknin, *et. al.* 2013). La mayoría de la gente está dispuesta a realizar pequeños sacrificios altruistas en situaciones propicias e incluso a favor de personas desconocidas y de manera anónima. Este hecho asegura que se trata de un altruismo incondicional sin expectativa de reciprocidad, ni directa ni indirecta (Frank, *et. al.* 1993b). En determinadas circunstancias críticas la gente es igualmente capaz de realizar actos de altruismo

heroico en beneficio de individuos desconocidos, como se pone de manifiesto en accidentes y catástrofes (Herranz 2013; Kretz 2011), si bien, el parentesco aumenta notablemente la propensión a cooperar y el grado de cooperación en tales situaciones (Burnstein, *et. al.* 1994). Lo que estamos planteando es que existen bases teóricas y empíricas suficientes para proclamar no sólo la cooperación condicional, sino además una base relevante de altruismo exoparental.

La psicología y la economía experimentales muestran que la especie humana no sigue al pie de la letra los patrones de comportamiento que pronostican las disciplinas académicas inspiradas en la teoría de la elección racional (Hardy-Vallée 2007: 941). Resultados parecidos se acreditan en las investigaciones de psicología cognitiva y neurociencia cognitiva social (Gneezy, *et. al.* 2012; Aknin, *et. al.* 2011; Dunn, *et. al.* 2008; Harbaugh, *et. al.* 2007; Moll, *et. al.* 2006). Estos hallazgos invitan a plantear hipotéticamente que una mayoría transcultural de individuos «normalmente manifiestan una clara tendencia a cooperar y a valorar la equidad» (Hardy-Vallée 2007: 942).

El modelo de evolución de la cooperación basado en el deme estructurado de interacción densa es una posibilidad teórica con valor explicativo, pero no la única. Los grupos humanos empíricos de la vida real tienden a ser de mayor tamaño y complejidad que los demes estables de reducida dimensión en los que resulta teóricamente viable la cooperación condicional²²⁷. La estrategia de cooperación real parece admitir un grupo óptimo de extensión superior al grupo de parentesco y al enclave social viscoso (Zeggelink, *et. al.* 2000). Está constatado que la historia humana en todas las culturas viene caracterizada por ser una historia de la acción colectiva, con grados variables de éxito, de una amplitud demográfica con frecuencia superior al número de Dunbar (Ostrom 2011[1990]). Existen igualmente pruebas que acreditan que la especie humana ha interactuado en modelos descentralizados de relación esporádicos y con desconocidos, desde tiempos que se remontan a un pasado evolutivo de entre 75.000 y 100.000 años (Gintis, *et. al.* 2008: 251-252; Smith 1998: 4). La movilidad de las poblaciones humanas, el intercambio comercial y genético-cultural, son algunos de los

²²⁷ Las características fundamentales de estas pequeñas agrupaciones (biológicas y culturales) de interacción densa son: una estructura social cerrada, homogénea e igualitaria, así como un tamaño poblacional limitado a 100-150 individuos, que no supera en ningún caso los 500 - 600 (Boehm 1993; Kosse 1990).

rasgos distintivos más relevantes de la especie, junto con el grupalismo altamente comprometido con la propia banda de cazadores-recolectores. Las relaciones exogámicas de cooperación han sido y son una característica de la especie benéfica para su supervivencia:

Si las condiciones en que los humanos emergieron son similares a las de los primates modernos y/o las sociedades contemporáneas de cazadores-recolectores, se puede reforzar el argumento de que hubo un flujo constante de individuos dentro y fuera de los grupos en esas sociedades. La exogamia en sí misma, de acuerdo a la cual los jóvenes machos o hembras se trasladan a otros grupos en busca de pareja, da lugar a una muy considerable mezcla intergrupala, y frecuentes encuentros con desconocidos y otros agentes con quienes la probabilidad de volver a interactuar en el futuro es muy escasa. (Gintis, *et. al.* 2005: 26).

En estas circunstancias los individuos están interactuando en un ambiente social diferente enclave relacional viscoso analizado en el capítulo previo²²⁸. Tampoco se trata, obviamente, del enjambre u hormiguero de insectos haplodiploides explicable de acuerdo al modelo de selección de parentesco. Aunque cooperar con desconocidos en juegos de tirada única, que es un caso de altruismo, no encaja en la concepción individualista y utilitarista de la psicología humana, sin embargo la evidencia empírica corrobora que esta anomalía teórica aparece en la ontogénesis antes de los tres años de edad (Warneken y Tomasello 2006). Se trata además de un comportamiento frecuente en la población adulta humana (Colman 2003; Camerer y Thaler 1995; Dawes y Thaler 1988).

Puesto que resulta evidente que los individuos humanos generalmente se hallan bien dispuestos a la cooperación recíproca e incluso en ocasiones al altruismo fuerte en beneficio de desconocidos, algunos trabajos propugnan que el altruismo no nepótico consiste en una exadaptación de los mecanismos, aún operativos, que gestionan las conductas altruistas de la selección de parentesco (Alexander 1974)²²⁹. Pedersen (2004) explica que el mismo sistema neuroendocrino que activa las emociones adaptadas para el cuidado de la prole y la vinculación con los parientes consanguíneos permitió a los

²²⁸ Téngase en cuenta que las civilizaciones post-neolíticas se caracterizan por generar ambientes urbanos comparativamente superpoblados en comparación con las de las eras pre-neolíticas. La diferencia es claramente relevante: de los 100-600 individuos en este último caso al millón largo de habitantes que poblaron la ciudad de Roma en el s. II dC., durante el gobierno de Trajano.

²²⁹ Por *exadaptación* se entiende una readaptación evolutiva de la morfología y del comportamiento de la especie hacia una(s) nueva(s) función(es) (Gould 1991; Gould y Vrba 1982). El término aplicado a la paleontología es el de *exaptación*. Darwin utilizó el vocablo “preadaptación”. En todos los casos alude a formas de innovación evolutiva que favorecen la eficacia del individuo que las desarrolla.

primeros humanos sentir afectos prosociales, que producen vínculos y comportamientos cooperativos con los miembros de otras familias de la tribu y, llegado el caso, con individuos desconocidos. Esta teoría, aún en fase exploratoria, cuenta con una creciente aceptación en psicología social (p.ej. Feeney y Collins 2001; Mikulincer y Shaver 2001). Las pruebas al respecto son evidentes en la psicología del desarrollo y la etología comparada (Taylor 2002: 275n₁₂).

Sobre esta base teórica en el presente capítulo descartamos la propuesta [1], la selección de parentesco (según se especifica en la Introducción del capítulo previo a este), como único factor explicativo concluyente del comportamiento altruista. En el apartado 2 se presenta la hipótesis central del capítulo: una naturaleza humana adaptada para la cooperación intraespecífica, o hipótesis del *homo cooperatibus*. En el apartado 3 se muestran evidencias acerca del comportamiento altruista exodémico y exogámico en especies de animales no humanos, así como evidencias de este tipo de comportamientos desde tiempos remotos en la evolución humana.

En consonancia con los principios del evolucionismo darwinista clásico, si el altruismo existe se debe a que es una solución adaptativa a los problemas de la vida humana. Esto quiere decir que este rasgo forma parte de una racionalidad ecológica, ya que de no favorecer la eficacia de los organismos que lo portan éstos serían “obligados” por las fuerzas de la evolución a “desprenderse” de él. La muestra fehaciente de un comportamiento disfuncional es su desaparición evolutiva (una no-replicación diferencial) por inadaptación biológica y cultural²³⁰. El paradigma clásico de la selección grupal no es la única alternativa al de la selección de parentesco en la explicación del comportamiento altruista. Así, en el apartado 4 se demuestra que el altruismo no sólo beneficia evolutivamente al receptor de los actos altruistas, sino que el individuo altruista también puede resultar beneficiado. En este apartado se explican dos factores de evolución del altruismo que no requieren para su éxito una selección de grupo: las adaptaciones referidas al refuerzo del comportamiento prosocial a través de

²³⁰ Valga como ejemplo del proceso de selección natural el cálculo consistente en que un pájaro mejor capacitado para el vuelo cuya configuración de alas le aporte una ventaja medida en probabilidad de reproducirse de 0,9999 en comparación con otro pájaro de la misma especie dotado con unas alas ligeramente peor adaptadas para la supervivencia y con una probabilidad de reproducirse de 0,9998, sus descendientes estarán 22.000 veces más capacitados para sobrevivir después de 100.000 generaciones (Magee 2006: 256).

una homeostasis individual mejorada, y aquellas otras relativas a la reciprocidad positiva activada por los sentimientos de simpatía, admiración y gratitud, así como a la reciprocidad social enmarcada institucionalmente en los sistemas culturales de prestigio.

En los apartados 5 y 6 se exponen dos hipótesis sobre la causalidad última del comportamiento altruista que justificarían las adaptaciones psicológicas prosociales en la especie humana. Tales hipótesis son:

- a) La interdependencia, o teoría del destino común (que se corresponde con la solución evolutiva [3] especificada según Hirshleifer, en la introducción del capítulo previo),
- b) y el papel que juega la dinámica poblacional, o perspectiva poblacional, en la eficacia de los individuos bien adaptados para cooperar selectivamente con otros cooperadores.

Tanto la teoría del destino común como la perspectiva poblacional aportan argumentos suficientes para explicar la existencia de motivaciones y preferencias universales de tipo altruista, así como la expresión epigenética de las mismas en la cultura e instituciones particulares de cada grupo o sociedad. Pasamos a analizar cómo operan estas causas últimas en la configuración de una psicología adaptada para el comportamiento altruista.

4.2.- Nacidos para cooperar

En su obra *Animales racionales y dependientes*, Alasdair MacIntyre señala un aspecto clave de la naturaleza humana que en su opinión se ha estudiado insuficientemente a la hora de explicar su extraordinaria sociabilidad. Se trata de la dependencia de los demás que existe siempre en todos los seres humanos y de manera especialmente intensa al menos en dos circunstancias prolongadas de la vida: inevitablemente en la infancia y la vejez, sobre todo, y además en los impredecibles quebrantos de la enfermedad y lesiones pasajeras, o en la discapacidad permanente. Los humanos se necesitan unos a otros para sobrevivir en el sentido más básico, que incluye

el desarrollo, reproducción y disfrute de calidad de vida. La vulnerabilidad y las aflicciones, y la dependencia de unas personas respecto a otras, son hechos «tan importantes por sí mismos que ningún autor que aspire a dar una explicación convincente de la condición humana puede dejar de concederles un lugar privilegiado» (MacIntyre 2001b: 15).

Como ya señalaron Richard Alexander y antes el propio Darwin, dado que la humanidad comparte aspiraciones y necesidades en un rango bastante homogéneo, la competencia intraespecífica y los procesos de selección individual tienen su lugar en la asignación de bienes individualizables y excluyentes tales como la comida, la vivienda, ropa, pareja sexual, etc. Sin embargo, los demás individuos son igualmente casi la única fuente de asistencia, apoyo en las dificultades, compañerismo en el logro de tareas complejas, protección y cuidados, además de fuente de conocimientos cuya difusión a través del aprendizaje ha posibilitado el rasgo cultural prominente de la especie. Existe un “interés” evolutivo en que las relaciones humanas se encaucen por una buena entente de coexistencia y cooperación. La selección individual mejora cuando los individuos se preocupan por los demás en aquéllos elementos de sus relaciones con ellos que repercuten en el aumento de su eficacia. Como sostiene Wilkinson (2001: 19),

... la calidad de nuestras relaciones sociales ha sido siempre vital para nuestro bienestar material. Dependiendo de nuestras relaciones con los demás, podríamos ganarlo o perderlo todo.

Consecuentemente, las relaciones sociales armoniosas y cooperativas son una variable explicativa de la eficacia, y la coevolución ha favorecido las pretensiones grupalistas que motivan para el logro de una sociedad amistosa, es decir, el instinto social de comunidad o instinto tribal. Por ello, y avanzando el argumento citando nuevamente a Wilkinson (2001: 93):

La cohesión social va asociada a la salud porque es un reflejo de lo tensos o relajados que son los contactos sociales, y de la ansiedad social subyacente (...) el decir “hola” a los vecinos y pertenecer a grupos locales está relacionado con la salud porque son indicativos de bajos niveles de ansiedad.

Según MacIntyre la conducta frente a esta dependencia real, o posibilidad de dependencia en una fecha indeterminada, es la reciprocidad; una reciprocidad cuyos códigos de refuerzo son las “virtudes”, y cuya expresión institucional es el entramado de normas formales e informales que regula la vida en sociedad y sus estructuras de solidaridad²³¹. Lejos de confiar al raciocinio y sus construcciones la articulación del extenso y complejo orden cooperativo de las sociedades humanas, MacIntyre insiste en que la vida moral humana «tiene como punto de partida la condición animal originaria del ser humano» (MacIntyre 2001b: 10). Siendo necesaria la reciprocidad cooperativa no sólo para vivir una vida acomodada, sino básicamente para sobrevivir, su origen hay que buscarlo en la evolución biológica, y su presencia en los innatismos prosociales. Curiosamente, la especie humana se la ha denominado *Homo sapiens*, cuando bien podría haber sido *Homo cooperatibus*. Destaca Ridley (1996: 249) en este sentido que: «la tendencia instintiva a la cooperación es la auténtica señal de identidad de la humanidad, y lo que nos diferencia de los demás animales». La cooperación intensa con los semejantes sin ser un animal haplodiploide es su rasgo más destacado dentro del reino animal, junto con la disposición de una psicología que le permite concebirse a sí mismo como sujeto. Examinemos la naturaleza y el origen de estos rasgos conductuales.

4.3.- Cooperación más allá de los lazos relacionales y grupos genealógicos

La cooperación y el altruismo fuera de los ámbitos familiar y de relaciones de proximidad resultan bastante infrecuentes en grupos de primates no humanos. Sin embargo, uno de los rasgos característicos de la especie humana, sobre cuya explicación evolutiva aún no se ha producido un consenso académico, es la cooperación incondicional con desconocidos²³². Aunque la teoría sociobiológica estándar se basa en la selección inclusiva de identidades genéticas, existen alternativas teóricas para la explicación del altruismo en las que la similaridad genética no es la única explicación (Fletcher y Doebeli 2009). Esto se constata incluso en interacciones de única jugada, por lo que tales actos pueden ser catalogados de altruismo fuerte (Janssen 2008; Cooper, *et. al.* 1996). Los humanos cooperan frecuentemente con extraños, y lo hacen en interacciones de aproximación en las que no se ha desarrollado suficiente confianza y

²³¹ El autor se está refiriendo a las normas morales de la cultura.

²³² Véase Silk y Boyd (2010: 229-230, y 240).

reconocimiento de replicadores comunes como para justificar un altruismo fuerte (Yamagishi, *et. al.* 2005). Estamos pues ante la incógnita del altruismo no nepótico y fuera del deme estructurado de la reciprocidad.

El enclave axelrodiano estricto, o deme estructurado autárquico con cinturón protector, tal vez nunca existió en estado puro como *única* forma de vida social humana. La humana no es una especie cooperativa porque haya evolucionado a través de un único mecanismo de selección, ya sea esta la selección de parentesco *à la* Hamilton, el altruismo recíproco *à la* Trivers, o alguna de las alternativas basadas en la tesis de la selección de grupo. La particular eusocialidad humana, que no se adapta específicamente a ninguna de estas explicaciones está fundamentada, además, por complejos sistemas relacionales de mutualismo vinculados a una reciprocidad inclusiva:

... existen buenas razones para creer que los primeros humanos se implicaron activamente en interacciones altamente relevantes para su eficacia fuera de las redes de parientes, y que muchas de esas interacciones fueron de una duración suficientemente corta como para que sólo una cooperación insólitamente beneficiosa evolucionara sobre la base de la reciprocidad. (Bowles y Gintis 2011: 226-227).

Los estudios de diferenciación genética en las etnias humanas revelan que los enclaves de cazadores y recolectores del Pleistoceno debieron interactuar con frecuencia y no sólo para el conflicto por los recursos, sino quizás con más naturalidad para intercambiar bienes, conocimientos y uniones sexuales (Bowles y Gintis 2011: cap. 7). La distancia geográfica entre las tribus muestra una deficiente correlación respecto a la distancia genética, lo cual sugiere que hubieron de producirse contactos de todo tipo entre clanes bastante alejados espacial y culturalmente, así como frecuentes y amistosos flujos y fusiones migratorias²³³. Sólo los grupos aislados durante largos períodos de tiempo por razones climáticas o geográficas muestran una significativa diferenciación genética poblacional. Los demes ancestrales no fueron pues grupos reducidos, estacionarios y endogámicos, ni biológica ni culturalmente aislados por murallas o cinturones protectores²³⁴. Este hecho puede explicar la insólitamente reducida

²³³ Un estudio reciente de genética de poblaciones que se remonta a los últimos 4.000 años es el de Hellenthal, *et. al.* (2014).

²³⁴ Estas vendrían a ser las características técnicas (estructura espacial, y finitud del tamaño de la población) de una red de interacción densa que, al menos teóricamente, garantiza la evolución de la cooperación (Taylor y Day 2004).

variabilidad genética de la especie humana si se la compara con la del genoma de otras especies de grandes simios estudiadas, tales como todas las de chimpancés, orangutanes, bonobos y gorilas (excepto la del gorila de montaña, no estudiada aún) (Prado-Martínez, *et. al.* 2013). El rasgo genético distintivo de la humanidad de lo humano es sin duda esta reducidísima diversidad genética intraespecífica, un fenómeno que suele ser el resultado de procesos evolutivos de cuello de botella, como los que ocurren cuando se produce un descenso rápido y drástico de la cantidad de individuos de una especie. En especies sociales tal homogeneidad genética puede acentuarse si los individuos supervivientes se reagrupan en una fase previa a su posterior diseminación geográfica. Esta teoría, coherente con los conocimientos disponibles sobre historia natural, puede ser explicativa del carácter eusocial de la especie humana, sin ser ésta especie haplodiploide, y explicaría en sí misma la tendencia propia de la humanidad de lo humano a un altruismo basal generalizado entre los miembros de la especie.

En otro orden de cosas, la teoría estándar del parentesco genético está siendo sometida a revisión, y el debate se encuentra abierto²³⁵. En efecto, en un mismo número de la revista *Science* varios autores han puesto en tela de juicio la relevancia de la semejanza genética en la explicación empírica del altruismo. Dickinson y Kienig (2003) cuestionan la relación de causalidad convencionalmente asumida entre parentesco y cooperación, y en sus estudios sobre las poblaciones de aves (animales que suelen agruparse y cooperar intensamente con parientes) proponen que la causalidad entre parentesco y altruismo no está mediada necesariamente y sólo por la homología (existencia de un ancestro común) sino por otra vertiente de la semejanza, que es la analogía (semejanza fenotípica sin parentesco genético común). Esto permite plantear que el altruismo con los parientes (p.ej. de progenitores hacia la progenie) es más una cuestión de “familiaridad” que de parentesco genético. El altruismo parental sería pues un efecto de la familiaridad, entendida como cercanía tolerante y no agresiva de la que surge un vínculo que motiva al altruismo, y no el linaje genético una causa de tal vínculo. El altruismo entre parientes puede, de este modo, derivarse de un apiñamiento

²³⁵ Véase Dugatkin (2007: 261-263). Diversas investigaciones han apuntado recientemente en esta línea, que busca el origen de la extrema sociabilidad humana fuera del parentesco y de la reciprocidad del enclave amurallado (Wilson y Hölldobler 2005; Nowak, *et. al.* 2010; Wilson 2012; García Lázaro, *et. al.* 2012). Un antecedente de la teoría de la sociedad aplicada a la evolución de la cooperación es Tullock (1990).

espacial reforzado y enriquecido, una especie de simbiosis por cercanía, independiente de los genes comunes.

Sinervo y Clobert (2003) han llegado a una conclusión parecida en sus experimentos con reptiles. Los lagartos, por ejemplo, se distribuyen en el territorio para protegerlo de intrusos, y en esta tarea no todos y no siempre buscan la compañía de sus parientes genéticos, sino que con frecuencia se asocian con individuos entre los que no se da una relación de parentesco. Ello no implica necesariamente que las vinculaciones se produzcan entre individuos con coeficientes de relación genética bajos: la semejanza genética viene explicada genéricamente por la homología biológica, y no sólo por la consanguinidad parental.

Los experimentos con dispersión limitada también han sido llevados a cabo por West, *et. al.* (2002). Estos autores han analizado la relación existente, de cooperación o competencia, entre individuos apiñados en un territorio. La conclusión es que esa cercanía puede dar pie a comportamientos tanto de cooperación como de competencia incluso entre parientes consanguíneos. La competencia entre parientes puede, de este modo, reducir o eliminar el mecanismo de la selección inclusiva basado en la genética común. De las relaciones entre no-parientes puede surgir la cooperación, e incluso el altruismo, si existen posibilidades de apareamiento fuera del linaje. Esto último conlleva asumir que la dispersión limitada puede incrementar el parentesco entre competidores genéticos, propiciando la cooperación, y reducirla entre parientes consanguíneos si los medios para la reproducción de los individuos son limitados²³⁶. La conclusión consiste en que en los apiñamientos surge una tensión entre dos fuerzas opuestas, que son la cooperación y la competencia. La competencia es más probable que surja y acabe dominando donde se da una dispersión asimétrica de los sexos de los individuos: en especies en las que pocos individuos de un sexo monopolizan a los individuos del sexo opuesto, o en las que uno de los sexos se dispersa (p.ej. por migración) más que el otro, resulta verosímil que surjan relaciones de competencia incluso entre hermanos gemelos. La cooperación sería pues una función del equilibrio de reparto en el acceso a los

²³⁶ Se apreciará mayor o menor cooperación respecto a la prevista teóricamente en los modelos de selección de parentesco cuando en la interdependencia social aparezcan complementariedades o incompatibilidades. Véase el modelo de Alger y Weibull (2012).

individuos del sexo opuesto. Las posibilidades de reproducción sexual entre no-parientes pueden, a su vez, ser la causa de una intensa cooperación dentro de enclaves en los que se apiñan individuos no-emparentados.

Este cuestionamiento teórico del altruismo basado en la replicación de identidades genéticas también ha sido presentado por autores del campo de la economía y en experimentación con individuos humanos [Fehr y Fischbacher (2003); Fehr y Gächter (2002)]. La pregunta implícita que se plantea sigue siendo la misma: ¿cómo ha podido evolucionar la cooperación fuerte entre individuos no emparentados?

Desde el punto de vista evolutivo la cuestión relevante para explicar el altruismo con desconocidos consiste en explicar qué estrategias despliegan los individuos para que sus acciones no debiliten su eficacia. Como indica Price (2008: 239), la clave para encontrar una salida a este problema teórico consiste en la integración de las teorías de la cooperación por selección individual dentro de las tradiciones del altruismo fuerte y débil, es decir, la selección de parentesco y la reciprocidad. La garantía de que el altruismo pueda evolucionar estriba en que los individuos practiquen conductas que incrementen su eficacia o la eficacia de los rasgos que sustentan el comportamiento altruista. La pieza a encajar en el modelo es el papel que juega la cultura en el funcionamiento del sistema de cooperación inclusiva humana, que se manifiesta empíricamente en la viabilidad de la acción colectiva:

De esta manera, los cooperadores pueden eludir ser explotados hasta la extinción por parte de *free-riders* que contribuyen menos que ellos, y la cooperación puede evolucionar vía adaptación a nivel individual. (Price 2008: 239).

La teoría convencional de la selección de grupo queda por ahora descartada, al demostrarse que la cooperación puede de este modo evolucionar por selección individual si la evolución genera paralelamente propensiones discriminativas, capacidades de integración en entornos relacionales y culturales, así como para la detección de cooperadores y la defensa y disuasión contra los defectores sistemáticos, tanto en interacciones diádicas como en complejos entramados de acción colectiva. La cooperación y el altruismo son dos estrategias de supervivencia capaces de incrementar la eficacia exponencialmente, siempre que el individuo cuente con las capacidades y

cauteladas necesarias para no ser explotado ni depredado en ambientes relacionales inciertos. Una expresión de la racionalidad ecológica de la especie humana es su psicología cooperativa inherentemente cultural. De ello que los humanos sean organismos, como ya hemos explicado, nacidos para cooperar, y por eso adaptados para extraer eficacia de la prosocialidad.

4.4.- El premio a la prosocialidad

Los humanos tienden instintivamente a dar y recibir información, cuidados, atenciones y regalos en general, a ser cooperativos tanto en relaciones estables como inestables, con conocidos y desconocidos, de tipo homófilo o heterófilo (Molina, *et. al.* 2008: 69). Sin embargo, que las conductas prosociales formen parte del etograma de la especie no significa que la prosocialidad empírica esté completamente codificada. En animales sociales y además culturales como el *Homo sapiens* el ambiente es esencial en la ontogénesis de la gama de comportamientos. Así, la exposición reiterada a experiencias de cooperación, empatía y compasión (*compassion training*), tanto con parientes, amigos e incluso con personas desconocidas y en interacciones sin posibilidad de reciprocidad, produce efectos prolongados de auto-refuerzo y refuerzo social (*behavioral plasticity*). El instinto admite refuerzos igualmente instintivos, que se activan o no en función de la experiencia que comienza en la vida prenatal²³⁷. La experiencia resulta determinante a efectos prácticos de la conducta. Los estímulos emocionalmente competentes para activar sentimientos morales y emociones prosociales se graban en la memoria emocional del individuo, transformando la operativa emotivo-cognitiva del cerebro y el patrón de toma de decisiones con efectos a largo plazo (Leiberg, *et. al.* 2011). Se ha constatado que las vivencias personales y el aprendizaje cultural adquirido tienen efectos a largo plazo, e incluso permanentes, en la arquitectura neurológica de la corteza parietal inferior, corteza dorsolateral prefrontal y la conectividad de esta última con el núcleo accumbens. Estas áreas del sistema nervioso se hallan directamente implicadas en los procesos psicológicos de la empatía e intersubjetividad, la regulación de las emociones y la producción de sentimientos

²³⁷ Se ha demostrado que los adolescentes y jóvenes que no han recibido cariño y cuidados parentales tienen una mayor probabilidad de incurrir en conductas y adicciones peligrosas para la salud. Existe una epigenética de la cooperación que requiere una estimulación temprana, incluso prenatal, de modo que el ambiente no cooperativo influye en la transmisión de tendencias antisociales dentro de los linajes, y con ello ciertos riesgos y problemas de salud (Wilkinson, 2001: 77-78).

agradables de refuerzo conductual (Weng, *et. al.* 2013).

Los rasgos de comportamiento con valor adaptativo de cualquier organismo están reforzados a través de innatismos propicios para su replicación. En los organismos dotados de cerebro y sistema neuroendocrino de recompensa este refuerzo puede consistir en experiencias psicológicas agradables, o incluso en mejoras en las capacidades generales de supervivencia y reproducción. Cuando los individuos humanos sostienen relaciones sociales que aportan eficacia, tales como la cooperación recíproca o conductas incluso netamente altruistas, la experiencia asociada viene a ser una sensación de bienestar, alegría de ánimo y mejora general del estado de la homeostasis (Aknin, *et. al.* 2013; Moll, *et. al.* 2006; Rilling, *et. al.* 2002). Esto puede justificar que a diferencia de otras especies de primates, como los chimpancés, en la especie humana se aprecia una motivación espontánea hacia las conductas prosociales. Asimismo existe evidencia de la valoración favorable de los individuos prosociales, una consecuencia de la preferencia universal por la cooperación y la búsqueda de cooperadores y posibilidades fácticas de cooperar (Tomasello y Vaish, 2013: 241). La evolución parece haber dotado a la especie humana de adaptaciones cognitivas, emotivas y motivacionales de modo que, citando la intuición filosófica de Spinoza:

La felicidad no es un premio que se otorga a la virtud, sino que es la virtud misma, y no gozamos de ella porque reprimamos nuestras concupiscencias, sino que, al contrario, podemos reprimir nuestras concupiscencias porque gozamos de ella. (Spinoza 2004[1677]: V, xlii; cursivas en el original).

La felicidad que menciona Spinoza consiste en que la propia biología evolucionada del organismo produce una gratificación psíquica a la cooperación, el altruismo y la honestidad que, parafraseándole nuevamente, supone un premio endógeno a tales comportamientos prosociales. No obstante, el disfrute de dichas recompensas *a posteriori* no tiene por qué ser la motivación que impulsa a actuar (Batson 1992: 70).

En el aspecto cultural, tampoco el aprendizaje conformista de los códigos de valor y normas sociales es una explicación suficiente para la ejecución de conductas altruistas. En efecto, en coherencia con Spinoza los comportamientos pueden ser culturalmente valorados por el premio natural que les sucede, pero no se llevan a cabo

necesariamente en el anhelo de esa valoración moral, sino que vienen a ser considerados virtuosos (valiosos desde el punto de vista reputacional) porque quien los realiza está demostrando ser capaz de sacrificar por los demás su eficacia personal, y eso es un signo de valía y distinción social. La cultura y la educación moral tienen capacidad de influir en los valores y la reputación por sí mismas, aunque es más verosímil que lo hagan si se acoplan bien con el etograma y aportan eficacia a quienes así obran (Boyd y Richerson 2002)²³⁸.

Hoy día comenzamos a saber acerca de la adaptación evolutiva consistente en que ese sacrificio prosocial a corto plazo de eficacia está correlacionado con un beneficio de orden mayor (para el grupo o incluso la especie), que incluye la posibilidad de un beneficio propio también superior a largo plazo. Fredrickson, *et. al.* (2013) han demostrado que a diferencia de la sensación de bienestar consecuencia de llevar a cabo “un noble propósito” (bienestar eudemónico), como es un moderado servicio altruista a los demás, la mera auto-gratificación hedónica a corto plazo (bienestar hedónico), como la producida por comida o sexo abundantes, no mejora el estado general de salud del individuo sino que incluso puede empeorarlo. Esto se debe a que el bienestar no hedonista influye favorablemente en el patrón de expresión de los genes en las células inmunes del organismo, aumentando y mejorando su operativa. El bienestar hedónico, por el contrario, se asocia a un aumento significativo en el perfil de expresión génica asociado al sistema del estrés²³⁹.

La sociabilidad tiene una importancia capital a la hora de explicar la cooperación y las relaciones de intercambio dentro de grupos, y constituye un instinto implantado en la evolución que aporta ventajas adaptativas. Particularmente, en la especie humana el aislamiento social, en sus diferentes manifestaciones físicas y psicológicas, es una de las causas no sólo de mortalidad del individuo, sino además, y más importante desde el punto de vista biológico, de ausencia de replicación genética y cultural (Holt-Lunstad, *et. al.* 2010; Fowler y Christakis 2008). La biología humana está especialmente

²³⁸ Esta apreciación choca con la teoría freudiana expuesta en *El malestar en la cultura*, así como con la tradición de pensamiento social y antropológico inspirada en la obra de Rousseau.

²³⁹ Se ha comprobado que la sobreactivación reiterada del sistema orgánico del estrés produce fallos en la replicación del ADN en los procesos de división celular, contribuyendo a la aparición de enfermedades y al envejecimiento prematuro (Buchan, *et. al.* 2011).

adaptada a la cooperación en grupos de diferente tamaño y complejidad, de modo que ésta contribuye a la producción de emociones y sentimientos agradables que se activan ante estímulos de cooperación y, a la vez, las capacidades biológicas de homeostasis y replicación funcionan mucho mejor (aparecen menos enfermedades, y por ello una superior eficacia reproductiva) si se practica frecuente y desinteresadamente la cooperación (Crimmins y Seeman 2004). Las dimensiones en las que los humanos practican las relaciones de apoyo social y ayuda mutua son muy diversas, si bien pueden clasificarse en tres patrones fundamentales (Molina, *et. al.* 2008: 62):

- a) De simpatía o apoyo *emocional* (relacionado con la experiencia de la comprensión intersubjetiva);
- b) apoyo *tangible* (que incluye ayudas directas, como alimento, donación de recursos, cuidados en la enfermedad, o facilitar medios de vida en general);
- c) y apoyo *informativo* (transmisión de técnicas, conocimientos, referencias y orientaciones de utilidad práctica).

El altruismo humano resulta motivado por causas próximas asociadas a la identificación vinculatoria con individuos concretos y marcadores abstractos de identidad cultural, así como un ferviente deseo de pertenencia a grupos parentales y culturales de diversa naturaleza. Esta innata “necesidad de pertenecer” (*need to belong*) es “una motivación poderosa, fundamental y extremadamente penetrante” (Baumeister y Leary, 1995: 497), que conduce a conformar y mantener relaciones interpersonales estables de apoyo mutuo. Además de esta propensión afiliativa, la principal emoción que motiva los comportamientos altruistas en general, tanto espontáneos y personalizados, como ritualizados en función de valores y normas sociales, es la emoción de la compasión (simpatía o piedad, utilizando otros términos sinónimos), que se define como “un estado emocional y motivacional caracterizado por sentimientos de calidez, amor y preocupación por el otro, así como deseo de ayudar a promover su bienestar” (Leiberg, *et. al.* 2011: 1). Los patrones de relación prosocial que surgen de ambos tipos de motivación influyen de manera determinante en la salud y el bienestar de los individuos, están asociados a procesos emotivos pre-rationales, se ven reforzados

por el afecto que surge del conocimiento e identificación mutuos, y desencadenan y condicionan procesos cognitivos, pautas emocionales y respuestas conductuales estereotipadas propias de la especie (Tomasello, *et. al.* 2012; Rand, *et. al.* 2012; Warneken y Tomasello 2006; Cole, *et. al.* 2007; Greene, *et. al.* 2003).

Herbert Simon (1983: 63) ha establecido tres condiciones explicativas de la evolución del comportamiento altruista relativas a la gratificación social de este tipo de conductas: (1) una tendencia psicológica a manifestar aprobación del comportamiento altruista observado en otros (y desaprobación del egoísmo); (2) una tendencia similar a responder con emociones de orgullo y vergüenza respectivamente a la aprobación y desaprobación de los otros; y (3) una tendencia general a recompensar el altruismo no sólo con aprobación simbólica sino también con oportunidades (o responsabilidades) que favorecen una eficacia diferencial. De este modo, podemos sostener que el premio a las conductas prosociales, como son la difusión altruista de conocimientos, descubrimientos e invenciones, es un premio endógeno y exógeno, ambos de refuerzo positivo, que la propia naturaleza humana manifiesta en la homeostasis orgánica, así como en la motivación para el comportamiento de gratitud diseminada a través de la interacción social. Se pueden plantear entonces dos factores causales de refuerzo a la prosocialidad cuya operativa es impremeditada:

- a) La endo-recompensa, o activación de procesos orgánicos que incrementan la eficacia desde el propio organismo y,
- b) la exo-recompensa, o activación de procesos psicosociales y culturales que incrementan la eficacia desde fuera del propio organismo.

4.4.1.- La endo-recompensa de una homeostasis mejorada

Las conductas cooperativas de ayuda mutua y altruismo, e interés por el bienestar de la comunidad (un bien con resultados de disfrute conjunto), son factores decisivos a la hora de explicar la eficacia biológica de los individuos, su bienestar y calidad de vida. Las personas que cuentan con un abanico más amplio y diverso de relaciones sociales de ayuda recíproca, que incluyen tanto las relaciones de fuerte

vinculación como aquellas otras de vínculo débil e incluso instrumentales, viven más años y tienen mejor salud.

La seguridad que aporta la disposición de relaciones de protección, ayuda y confianza en todas las edades de la vida, incide directamente en la gestión más saludable del mecanismo neuroendocrino del estrés (circuito hipotálamo-hipófiso-suprarrenal). Esto favorece una autorregulación hormonal benéfica por la menor sobrecarga por debilitamiento del sistema inmunitario. Como explica Taylor (2002: cap. 5) citando algunos de los estudios publicados por dos investigadores pioneros en la materia, Lisa Berkman y Leonard Syme, *cæteris paribus* se aprecia una significativa correlación entre la carencia de lazos comunitarios y la probabilidad de muerte. En otro estudio Wilkinson (2001: 32 y 57) asegura que las relaciones de dominación combinadas con el aislamiento social explican una diferencia en los índices de mortalidad entre 2 y 4 veces superior en aquellos individuos dominados y que carecen de un entramado social de apoyo. Esta diferencia en los índices es similar a la provocada por la exposición frecuente a sustancias tóxicas. Especialmente significativas para la vida son las relaciones de vinculación fuerte, como las de familia y amigos íntimos; si bien las relaciones comunitarias de vínculo débil, e incluso las relaciones con extraños en circunstancias puntuales (como las de amenaza compartida) también están demostradas como eficaces estrategias para la supervivencia (Diamond 2001).

El meta-análisis de Holt-Lunstad, *et. al.* (2010), realizado con una muestra de más de 300.000 individuos, concluye en que las relaciones amistosas de cooperación mutua aumentan la probabilidad media de supervivencia individual en un 50%, un beneficio en términos de salud y supervivencia similar al obtenido de abandonar el tabaquismo o el alcoholismo. Existe además una evidencia creciente de que quienes dedican parte de sus recursos a ayudar de forma desinteresada a otros, por ejemplo mediante acciones de voluntariado, gozan de mayor bienestar, mejor salud física y mental, y una inferior tasa de mortalidad que los individuos aislados y dedicados a objetivos exclusivamente egocéntricos. Esto constata que no sólo produce una mejora de la eficacia el recibir cuidados y ayudas de los demás, sino que incluso el realizarlos a favor de otros influye favorablemente en la eficacia del individuo.

En efecto, tanto en el estudio de Stephen G. Post (2009) como en el meta-análisis de Jenkinson, *et. al.* (2013), se presentan los antecedentes y una revisión de trabajos académicos de las áreas médica y psicológica, en los que se muestra la relación positiva existente entre el estilo de vida altruista y los beneficios experimentados en la salud, y a la postre en la esperanza de vida. Dichos beneficios no están necesariamente relacionados con la gratitud o la reciprocidad que puedan recibir los altruistas de sus beneficiarios o de terceros (aunque la reciprocidad incide positivamente en la motivación y en los beneficios experimentados por la persona altruista), sino con que el propio organismo del altruista gestiona más eficazmente su propia homeostasis. Parece existir una programación implantada en la naturaleza humana cuya funcionalidad consiste en promover el bienestar del grupo explicable por factores biosociales y culturales aún no bien conocidos. Así, por ejemplo, Post cita el estudio pionero llevado a cabo por el psicólogo David McClelland con los estudiantes de Harvard, en el que se demuestra cómo los estudiantes que visualizaron una película sobre el trabajo de Madre Teresa con huérfanos en Calcuta registraron un fortalecimiento inmunológico significativo en la saliva, en comparación con el efecto neutral que produjo la visualización de otra película por parte del grupo de control²⁴⁰. La misma composición inmunológica de S-IgA permaneció elevada una hora después de ver el documental en los estudiantes a los que se pidió que recordaran experiencias personales en que dieron o recibieron amor.

Una conclusión de gran importancia derivada de estas investigaciones es que los individuos que practican el apoyo interpersonal en redes sociales (p.ej. de amigos, parientes y vecinos) mejoran su salud y longevidad respecto a los que no lo hacen. Este beneficio resulta de una doble vía: la consecuencia de la propia homeostasis mejorada, y la del apoyo recíproco recibido de los demás integrantes de la red (Dunn, *et. al.* 2008). No obstante, el primer factor aislado es suficiente para explicar el efecto benéfico del altruismo, puesto que una vez establecido el hábito altruista se ha comprobado que recibir también apoyo de los demás no influye de manera relevante en la mortalidad (Brown, *et. al.* 2003).

²⁴⁰ El estudio aludido es el de McClelland y Kirchnit (1988).

El premio a la prosocialidad, en todo caso, resulta acrecentado desde un punto de vista relacional más que por un estilo de vida solamente de sacrificio y abnegación unidireccional. Así, el capital social de los individuos, su participación en relaciones multidireccionales de cooperación abundante y de buena calidad, aumentan comparativamente las posibilidades de una eficacia biológica y cultural enriquecidas. Se ha demostrado que las personas sometidas al odio, estigma, indiferencia o aislamiento social prolongados sufren alteraciones en la transcripción epigenética de las instrucciones de biorregulación y respuesta inmunológica de los genes responsables de la respuesta anti-inflamatoria, así como una respuesta desproporcionada de los genes que activan procesos pro-inflamatorios (Cole, *et. al.* 2007). La inadaptación característicamente humana a estados prolongados de soledad hace que los individuos sometidos a situaciones de ausencia de afectividad y lazos sociales, apoyo y relaciones, experimenten profundos estados psicológicos de soledad, que influyen en la depresión psíquica e inmunitaria. Esto deteriora la salud y la calidad de vida debido a una superior prevalencia de las enfermedades infecciosas, y por la sobreactivación del sistema del estrés y los efectos nocivos de emociones negativas como la tristeza y la desesperanza (Pressman, *et. al.* 2005).

4.4.2.- La exo-recompensa derivada de los sentimientos de admiración y gratitud, y los sistemas culturales de prestigio

Dentro de la gama de pulsiones de tipo prosocial existentes en la naturaleza humana figura la propia aversión a la antisocialidad. Antisocial es quien incumple las normas que regulan la vida social, y por esto activa emociones negativas y de rechazo en quienes sí las cumplen. La reputación del individuo está asociada a suscitar en los demás emociones positivas, por lo que las personas sin sociopatías se esfuerzan para conseguir ser acreedoras de esas emociones.

La causa próxima de la aversión al oprobio social se halla en dispositivos neurofisiológicos innatos, reforzables biológica y culturalmente, adaptados para alinear la conducta del individuo con las normas morales y sociales. La interiorización espontánea de normas de conducta obedece al instinto de conformismo (*docility*), que ha sido relacionado con el comportamiento altruista y los procesos de selección grupal

(Eibl-Eibesfeldt 1982; Henrich 2004; Hayek 1982[1976], Hayek 1990[1988]). Castro y Toro (2004) lo relacionan además con la adaptación humana a las normas sociales transmitidas culturalmente. De este modo, la asimilación de reglas generales de conducta, lejos de constituir solamente un proceso de selección grupal, aportaría eficacia al individuo al acomodar su comportamiento a entornos relacionales predecibles, de coordinación espontánea, y ricos en información, creencias y conocimientos compartidos (Simon 1990).

Ante un dilema decisorio entre cumplir o no las reglas socio-morales los individuos sin trastornos sociopáticos experimentan una sobreactivación espontánea del sistema límbico (la amígdala y la región circundante al hipocampo) con consecuencias psicológicas desagradables (Sommer, *et. al.* 2010). Este dispositivo autorregula el comportamiento cuando se enfrentan las motivaciones puramente hedónicas con aquellas que activan conductas de inclusión y cohesión social. La contrapartida de esta aversión a las manifestaciones de antisocialidad consiste en las espontáneas simpatías y complacencias intragrupales e intraespecíficas, que refuerzan las conductas prosociales a través de la interacción social. El instinto que contribuye a cohesionar los grupos mediante procesos psicológicos de identificación con estereotipos y agregados sociales consiste en una atracción por el sentido de pertenencia a grupos de diversa naturaleza. Richerson y Boyd (2001) lo denominan instinto tribal (*tribal instinct*).

Si las conductas altruistas forman parte del etograma humano, sólo aquellas excelsas y significadas serán valoradas reputacionalmente. Los individuos capaces de sacrificarse por otros fuera su linaje familiar demuestran estar dotados de capacidades extraordinarias, dado que pueden sobrevivir y replicarse en inferioridad de condiciones²⁴¹. Si el altruismo es una señal de calidad diferencial, su presencia en el *pool* genético de la especie resulta benéfica, dado que contribuye a propagar eficacia -de altruistas y no altruistas- mediante externalidades positivas. *Cæteris paribus*, sólo los individuos de mayor calidad atraerán a potenciales socios y compañeros sexuales más que la media, y aquéllos que sean capaces de reconocer su valía tratarán de favorecer su

²⁴¹ Millet y Dewitte (2007) han hallado evidencia experimental de que el altruismo es un buen indicador de eficacia congénita general, que se concreta en una superior inteligencia social, capacidad de liderazgo y mejor salud congénita.

supervivencia y replicación porque es adaptativo beneficiarse de su existencia.

La especie humana está adaptada para generar estados emotivos de exaltación, admiración y gratitud (*other-praising emotions*) ante las conductas que señalizan excelencia y calidad individual (Vianello, *et. al.* 2010; Algoe y Haidt 2009). Dada la propia naturaleza de estas emociones sociales, que motivan para premiar los comportamientos meritorios por encima de lo normal, existe del mismo modo una adaptación filogenética consistente en la motivación -no necesariamente patológica- de búsqueda de la alabanza de los demás. Esta propensión conduce, en determinados individuos, a la persecución esforzada del ensalzamiento y admiración social incluso mediante conductas que deterioran comparativamente su eficacia individual. De este modo, la honesta (no fraudulenta) prosocialidad extraordinaria, exitosamente señalizada, motiva el elogio y la simpatía de otros, lo cual contribuye a preservar este rasgo dentro de la evolución de la especie²⁴².

En otro orden de cosas, la calidad individual fidedignamente revelada, y la respuesta social a la misma, son transmitidas culturalmente y por lo tanto coevolucionan a través de sistemas institucionales de recompensa basados en el prestigio concedido (con frecuencia simbólico) y la dinámica de los comportamientos imitativos²⁴³. La acción de recompensar consiste en acciones de reciprocidad positiva, directa o indirecta, débil o fuerte, destinadas al altruista significado, así como la realización con él de alianzas que confieren ventajas selectivas recíprocas. Estas acciones de recompensa se pueden canalizar a través de sistemas socioculturales que institucionalizan relaciones de estatus, liderazgo y autoridad, los cuales otorgan al destinatario evidentes beneficios de replicación (biológica y cultural), ya que el prestigio activa la concesión de privilegios

²⁴² Sobre los modelos de selección de calidad mediante señalización altruista (*costly signaling as a handicap*) véase Magee (2006: 261-262). La señalización, que ha de ser lo suficientemente costosa y exclusiva para evitar ser objeto de fraude por parte de impostores, permite atraer a beneficiarios que puedan ser parejas sexuales o aliados de cooperación. Dos experimentos recientes con individuos humanos son los de Bereczkei, *et. al.* (2010) y Farrelly, *et. al.* (2007). El primer estudio concluye en que es el factor reputación y la atracción de cooperadores altamente comprometidos, y no tanto de parejas sexuales, lo que coadyuva a la evolución del altruismo de alta calidad. El segundo muestra efectos significativos de la señalización altruista en la selección sexual.

²⁴³ Téngase en cuenta que la exorrecompensa no necesariamente ha de ser onerosa para ser eficaz. Puede tratarse de una señal de complicidad, un elogio o el ensalzamiento público de una reputación (Ross 2008: 132). Sobre los modelos de replicación cultural por aprendizaje e imitación véanse Boyd y Richerson (2002), McElreath, *et. al.* (2003) y Börgers y Sarin (1997).

tales como el mando jerárquico, influencia y autoridad moral, deferencia en el acceso a recursos escasos u obligaciones generales, y difusión imitativa de sus conocimientos, creencias, valores y conductas (Henrich y Gil-White 2001).

La base de la emoción social de la gratitud se halla en una respuesta innata a la conducta de los altruistas benefactores, que activa la simpatía identificatoria con los sentimientos del receptor de la ayuda (sea uno mismo u otro), si bien con una evaluación cognitiva adicional que, en la especie humana, modula la emoción y el sentimiento de alegría empática (Lamm, *et. al.* 2007). Esta capacidad de evaluación cognitiva permite simpatizar con la simpatía de otros que sienten gratitud, y de ello es posible una gratitud vicaria que motiva las acciones de premio o castigo a terceros a través de la reciprocidad indirecta (Singer, *et. al.* 2006; De Quervain, *et. al.* 2004; Seymour, *et. al.* 2007). De ahí que la respuesta instintiva de reciprocidad puede ser en nombre propio o de otros, y la emoción tiene un impacto directo tan poderoso en la motivación que no se produce (del todo) un efecto expulsión (*crowding out*) de la deuda de gratitud por la existencia de recompensas ya realizadas por terceras personas (Andreoni 1993; Singer y Steinbeis 2009). La acción de reciprocidad se puede realizar independientemente de si dicha acción tiene o no un coste/beneficio para quien la realiza (Fehr, *et. al.* 2002; Fehr y Gächter 2002)²⁴⁴. Esto revela la existencia de un valor psicológico intrínseco en las acciones de gratitud.

Los fundamentos epigenéticos de las conductas de gratitud afloran a la cultura a través de conductas ritualizadas de premio a la excelencia y los valores, o lo que es lo mismo, el premio a la calidad (social y cultural) del individuo en cuestión. Pagar un premio a la calidad significa un coste aparentemente innecesario para quien lo paga por encima de lo que correspondería en una situación sin externalidades. Sin embargo, quien predispone a su favor a los individuos de mayor calidad tiene mayor probabilidad de acabar beneficiándose de algún modo de su presencia (*neighbour effect*), lo cual incluye la posibilidad de una provechosa alianza cooperativa con ellos, especialmente si unos y otros son cooperadores y son capaces de señalar eficazmente su semejanza (Gächter y Thöni 2005).

²⁴⁴ Lo cual no equivale a que la cuantía del coste/beneficio no influya en la modulación del comportamiento (Anderson y Putterman 2006).

Singer (2000: 79) expone dos ejemplos de expresiones ritualizadas de la gratitud social hacia los individuos que señalizan una calidad por encima de lo normal. Dicha expresiones son socialmente funcionales y contienen un ponderable valor de supervivencia:

Entre algunas tribus amerindias de las Grandes Llanuras, el guerrero puede hacer voto solemne de luchar hasta la muerte en la inminente batalla. Entonces se le permite, en los días que median hasta el combate hacer el amor con cuantas mujeres bien dispuestas él quiera. Cuando, durante la Primera Guerra Mundial, las chicas entregaban plumas blancas a los hombres en edad militar que iban de paisano, seguramente estaban haciendo algo similar, aun cuando las retribuciones reproductivas no fueran tan directas como en el caso anterior²⁴⁵.

En esta línea, López y Valdaliso (1999: 37-38) explican que los creadores culturales (científicos, artistas, tecnólogos o innovadores institucionales) son los grandes altruistas de la modernidad. El innovador de las sociedades actuales expresa sus instintos altruistas con creaciones puestas a disposición de la especie a través de la sociedad y sus instituciones. Éstas le otorgan recompensas culturales que no necesariamente consisten en una eficacia reproductiva (sexual) directa sino que, sin descartarla, más bien premian con una mejora indirecta de su eficacia a través de sistemas de honor y prestigio, y/o el beneficio de la explotación temporal de derechos exclusivos de propiedad intelectual (como patentes y derechos de autor). Ambos tipos de gratificación permiten igualmente un mejor acceso a la larga para tanto al altruista, como a sus familiares o su progenie cultural, a posibilidades de reproducción biológica diferenciales. El comportamiento innovador, de este modo, resulta una modalidad de comportamiento altruista adaptativa por medio del refuerzo cultural. Las innovaciones institucionales que fortalezcan este refuerzo, lógicamente, favorecerán la supervivencia de los innovadores altruistas y de ello la producción y difusión de innovaciones.

Así pues, la epigénesis de los instintos humanos y su reflejo en las instituciones sociales hacen posible la evolución del comportamiento prosocial en la naturaleza humana, de modo que ésta se halle equipada de procesos de causación próxima que motivan para llevar a cabo dichos comportamientos. Aun siendo altruistas, no obstante, la causación última que explica la filogénesis coevolutiva de tales comportamientos es

²⁴⁵ El acto simbólico de entregar la pluma blanca es un barato castigo que consiste en acusar elegantemente de “gallina” (cobarde).

compatible con la supervivencia de los mismos. Pasemos ahora a analizar otros procesos de causación explicativos de la filogénesis del comportamiento prosocial.

4.5.- La teoría del destino común

En sus estudios de los procesos de vinculación entre individuos humanos adultos Diamond (2001) ha hallado que existen fundamentos neurofisiológicos y neuroendocrinos que explican las conductas de ayuda no sólo con parientes consanguíneos, sino también con personas completamente desconocidas. Un factor determinante para la activación de tales dispositivos orgánicos es la percepción de una amenaza colectiva compartida y la existencia de procesos sociales de complementariedad (Hurlbert, *et. al.* 2000; Alger y Weibull 2012). En esas circunstancias, que pueden extrapolarse a situaciones y objetivos que requieren una acción conjunta y “sin fisuras”, la mayoría de los humanos tienden a ser cumplidores de las normas sociales y generosos con los demás²⁴⁶. Los humanos no sólo saben buscarse unos a otros cuando la interdependencia es técnicamente necesaria para el logro de metas individuales o colectivas, sino que además, y a diferencia de los chimpancés, buscan la colaboración aunque no sea del todo necesaria, y se sacrifican altruistamente por sus semejantes en circunstancias y modos que no se dan en las demás especies de primates²⁴⁷.

La interdependencia de los organismos se manifiesta en la existencia de un destino común que les hace necesariamente cooperativos. Así, la necesidad de cooperación, y por lo tanto el nivel de cooperación empírica C en la acción colectiva, es una función del *grado de comunidad* en los objetivos de los individuos que integran la población. Éste viene determinado por dos variables: el *coeficiente de convergencia* (c),

²⁴⁶ Según el modelo planteado por Fletcher y Doebeli (2009) para que el altruismo no parental pueda evolucionar es preciso que los individuos portadores del genotipo altruismo se vean *necesariamente* involucrados en acciones colectivas con individuos que expresan el fenotipo altruismo, siempre y cuando de su cooperación obtengan una eficacia neta promedio igual o superior a la de defectores y cooperadores condicionales.

²⁴⁷ Explican Tomasello y Vaish (2013: 235) que los chimpancés sólo cazan en grupo cuando no les es estrictamente necesario para sobrevivir. No lo hacen durante la estación seca, cuando el alimento vegetal es escaso, sino en la época de lluvias, cuando la fruta y los vegetales son más abundantes. Esta carencia de una innata actitud cooperativa, que motiva para la interdependencia y facilita su coordinación, es propia de un animal que revela preferencia por la consecución solitaria de objetivos, aun estando capacitado para actuar coordinadamente en procesos de acción conjunta. Los chimpancés parecen comportarse de acuerdo a las pautas descritas por Rosseau (1973[1754]: 73-74) en su famosa alegoría de la caza del ciervo.

que mide el grado de coincidencia en los objetivos comunes particulares; y la *tasa de cooperatividad esperada* (T), o compromiso esperado de los individuos, que mide el grado de cooperación de *todos* los integrantes de la población que es *necesario* para conseguir la parte común de los objetivos:

$$C = f(c, T) ,$$

estando c y T relacionadas positivamente con C .

La supervivencia puede depender de que los individuos confíen en la capacidad cooperativa de los demás, y que esto les motive para formar grupos de socios mutuamente comprometidos (Tomasello y Vaish 2013: 239). Este destino común puede consistir en una amenaza o beneficio común, individualizable o no, que afecta a todos por igual y ante la que no caben conductas oportunistas²⁴⁸. De practicarse el oportunismo se pondría en peligro la consecución de la meta colectiva en la que consiste el propio destino común. Éste asimismo puede ser una circunstancia que obliga a los individuos a estar unidos y a comportarse de acuerdo al bien o interés común, dado que no cabe otra posibilidad. El destino de estar unidos puede ilustrarse con el ejemplo de los hermanos siameses, o las personas que viajan en un mismo vehículo (un avión o un cohete espacial, por ejemplo) que no pueden abandonar en un momento dado²⁴⁹. La existencia de indivisibilidades acentúa la comunidad de objetivos y procedimientos, y de ello el fortalecimiento de la cohesión grupal. Cuando existe un destino común de este modo, los organismos, u otras unidades de selección, desarrollan comportamientos y actitudes prosociales, dado que las conductas conducentes al conflicto y la ausencia de unidad de propósito son incompatibles con la supervivencia del individuo que pudiera sentir motivación de aprovecharse de los demás (Oberholzer-Gee 2007). En el plano de las comunidades, la existencia de procesos de destino común como bien común contribuye a que la estructura de pagos de la acción colectiva se transforme, conduciendo a puntos focales de cooperación:

²⁴⁸ Dos ejemplos ilustrativos son la defensa conjunta frente a depredadores u otros peligros naturales, y la caza comunitaria de grandes animales.

²⁴⁹ Adviértase que el destino común de las unidades que viajan en un mismo vehículo que no pueden abandonar es una metáfora válida para representar el destino común de los genes que integran la macromolécula del ADN de un organismo. La única forma de “salida” del vehículo es la replicación, en el caso humano sexual, y por ello mismo establece un destino compartido.

La comunidad sólo puede “mantenerse unida” si los miembros saben que tienen una identidad de destino. En la práctica esto significa dos cosas. Los miembros deben identificarse con el pasado y (aún más importante) con el futuro de la comunidad, de tal modo que su propio futuro y el futuro de la comunidad sean uno y el mismo. Nos levantamos y caemos juntos. Esta es otra manera de decir que profesamos un compromiso *incondicional* con nuestra comunidad (...) La identidad o el destino también significa que los miembros se relacionan entre ellos en una atmósfera de confianza mutua (...) Si alguna persona reclamara que tiene un derecho inherente de ganar para sí misma a expensas de los demás, el tejido entero colapsa. (Shenker 1986: 93; cursivas en el original).

En un marco evolutivo esta restricción de supervivencia conjunta, como la amenaza compartida de un grupo étnico rival, que impele a actuar con unidad de propósito y lealtad al grupo en pos de la propia supervivencia, juega a favor de la evolución de innatismos prosociales, así como procesos orgánicos, psicosociales y culturales de refuerzo de los mismos (Bowles, *et. al.* 2004).

La perspectiva del destino común se manifiesta tanto en unidades de selección del microuniverso como del macrouniverso, y la regulación consiste en la evolución a través de la innovación de órdenes cooperativos, o «ley de cooperación natural» (Nowak 2006a: 1.563)²⁵⁰. Así, aunque en el neo-darwinismo que triunfó en las décadas de 1960 y 1980 fue aceptado el principio reduccionista de que la unidad de selección es el gen y no el organismo (Williams 1996, Dawkins 1993; Hull 1988), como si los genes compitieran entre ellos, sin embargo, los estudios posteriores han demostrado que los genes individuales se replican a sí mismos agrupándose con otros genes en estructuras ordenadas (el ADN), y que no pueden propagarse sin ser transportados por organismos contruidos por las propias estructuras ordenadas de genes agrupados. Dentro de estas macroestructuras relacionales de carácter cooperativo definitorias del fenómeno vida, los organismos se desarrollan en la interacción con el medio, y es en su expresión fenotípica en la interacción con el medio donde se produce la selección y evolución de los genotipos. Un gen, de este modo, “coopera” con otros genes para construir organismos individuales que posibilitan su propagación conjunta (Mayr 2000).

Obviamente, admitir que el gen necesita al organismo para propagarse conlleva que aun siendo el gen una entidad “egoísta” que puede perseguir “su” propio interés, y que puede entrar en conflicto con otros genes egoístas en conjunción con los cuales se

²⁵⁰ La evolución en sí misma es un continuo proceso constructivo de estructuras de cooperación (Maynard Smith y Szathmary 1999).

construye el organismo, sin embargo, la cooperación es necesaria, y la existencia de cooperación plantea una nueva perspectiva acerca de la unidad de selección: el individuo-gen *versus* la población-organismo (Sober 1996: cap. 4). Aunque los genes fueran egoístas no les quedaría más remedio que cooperar por su propio interés, dado que la alternativa consiste en una inverosímil replicación en soledad, o romper el equilibrio del grupo de genes con estrategias oportunistas que pueden dar al traste con la supervivencia del organismo, y por lo tanto con la del propio vehículo que coadyuva a replicar al gen junto con los demás genes²⁵¹. De este modo, si los genes individuales fueran como agentes según la teoría ortodoxa de la elección racional, optarían por la supervivencia y replicación dentro de grupos cooperativos en los que se cumplen las reglas de cooperación, en lugar de la competencia solitaria y desarraigada frente a otros genes (Frank 1987). Para replicarse con mayor éxito al gen no le queda más opción que interesarse por el destino de otros genes con los que coopera y a los que necesita.

Esta configuración relacional de destino(s) común(nes) se manifiesta en el macrouniverso de los organismos en la expresión de una actitud de respeto e interés por todo aquello que posibilita el logro de objetivos compartidos que constituyen el destino común. Al nivel del fenotipo y del grupo-población el destino común se erige en una causa última propiciatoria de la psicología adaptada para la cooperación y el altruismo. Hayek considera esta relación de fuerzas un orden espontáneo:

Al igual que los organismos biológicos, en las formaciones sociales espontáneas observamos con frecuencia que las partes se comportan o se mueven como si su objetivo fuera la preservación de los conjuntos. Nos encontramos una y otra vez con que *si* alguien deliberadamente tuviera como objetivo conservar la estructura de esos conjuntos, tendría que provocar precisamente esos movimientos que, en realidad, ya tienen lugar sin que nadie los dirija. (Hayek 2003[1952]: 132; cursivas en el original).

Desde otro punto de vista, se alude a la capacidad de establecer relaciones de sociedad y reducir así los riesgos inherentes a una existencia solitaria y errante:

... la vida cotidiana contiene pedacitos de altruismo por todas partes (...) Por ejemplo, cuando estoy en un avión, todo el mundo es agradable con los demás, aunque seguramente

²⁵¹ La selección actúa directamente sobre fenotipos -organismos- e indirectamente sobre genotipos -ADN, y genes ubicados en alelos específicos-. De este modo, la evolución es posible porque existen *organismos* que sobreviven y se reproducen con frecuencia diferencial respecto a otros. Véase Sober (1996: 155).

no volverán a verse nunca. ¿Por qué esa corrección? (...) (porque) Esto hace que funcione la sociedad. (Gintis 2006b; paréntesis añadido).

La preocupación por los intereses de los demás impulsa un ámbito relacional de mutualismo ventajoso para todas las partes intervinientes respecto a las alternativas asocial o antisocial:

Si una criatura pone un bien más allá de sus intereses individuales, es porque su destino se encuentra inextricablemente unido al de un grupo: comparte el destino grupal. La mejor esperanza de inmortalidad de una hormiga estéril es su reproducción vicaria a través de la prole de la reina, igual que la mejor esperanza de seguir viviendo que tiene el pasajero de un avión es a través de la supervivencia del piloto. (Ridley 1996: 39).

Se puede comprender de este modo la evolución de un *instinto de preocupación por el bien de la comunidad*, incluso desde una perspectiva motivacional original no necesariamente altruista (Piliavin y Charng 1990)²⁵². Frans de Waal ha denominado “preocupación por la comunidad” a este instinto de cooperación que incluye toda la gama de conductas cooperativas: desde la cooperación condicional, pasando por un altruismo basal con individuos desconocidos, hasta el altruismo fuerte propio de las conductas altamente comprometidas. Según De Waal cuando el individuo encuentra motivación para actuar en este sentido comunitario, «más bien, parece esforzarse por conseguir el tipo de comunidad que más pueda satisfacer sus intereses» (De Waal 1997: 265). Por lo tanto, una actitud eficaz para participar en acciones colectivas, evitando así tragedias sociales, no depende únicamente de la disposición de habilidades cognitivas desarrolladas, sino que es suficiente con un repertorio de instintos prosociales que propicien la autorregulación del grupo sin un Leviatán externo:

... la preocupación por la comunidad puede definirse como *el interés de cada individuo por fomentar aquellas características de la comunidad o del grupo que aumentan los beneficios que sacan ese individuo y su familia de vivir en dicha comunidad*. (...) los animales sociales han sido seleccionados para inhibir acciones capaces de mermar la armonía del grupo y para procurar conseguir un equilibrio óptimo entre la coexistencia pacífica y la satisfacción de los intereses particulares. Para la evolución de la preocupación por la comunidad, el hecho de que los animales se den cuenta del modo en que su conducta repercute en el grupo en general no es más importante que el que sean conscientes de la relación entre el sexo y la

²⁵² Como precisa Dawkins, entre genes egoístas es posible la defeción si ésta estrategia aporta más eficiencia que la cooperación en un momento dado:

Nuestros propios genes cooperan entre sí, no porque sean *nuestros*, sino porque comparten la misma salida –espermatozoide u óvulo– en el futuro. Si algún gen de algún organismo, como el humano, pudiera descubrir una manera de propagarse que no dependiera de la ruta convencional a través de espermatozoides u óvulos, la tomaría y sería menos cooperativo. (Dawkins 1993[1976]: 315).

reproducción para desarrollar unas estrategias de reproducción óptimas (...) Cuanto mayor es la conciencia social de una especie, más se darán cuenta sus miembros de que los acontecimientos a su alrededor rebotan por toda la comunidad hasta que aterrizan frente a su propia puerta. (De Waal 1997: 267-268; cursivas en el original).

Asimismo, otro instinto evolucionado para reforzar la cohesión social en pos del cumplimiento de un destino común es la vinculación a personas, grupos o estructuras ideológicas culturales, que resulta motivada por identificación y sentimientos de pertenencia, lealtad y amor (Eibl-Eibesfeldt 1996: cap. 4). El conocimiento de una amenaza o de un posible beneficio compartidos que afectan al grupo con el cual el individuo se identifica activa el llamado por Lorenz (1978[1963]: 302-304) instinto de entusiasmo militante. Esos sentimientos endogámicos (como la identificación, pertenencia, afectividad y lealtad) tienen un claro origen adaptativo, dado que contribuyen a motivar comportamientos que incrementan la eficacia (Fiske 2002: 171; Pinker 2004: 514 y 519).

En efecto, la evolución ha desarrollado un cerebro humano dotado de dispositivos neurofisiológicos capaces de producir experiencias sentimentales agradables tales como cariño, vínculo, apego, simpatía, compañerismo, confianza, cercanía, afinidad, etc., que motivan al individuo y lo comprometen a ayudar a los demás miembros de un grupo en el que, o bien *todos* se ayudan unos a otros en cuestiones y momentos puntuales, o bien *todos* se encontrarán en una situación peor si no lo hacen (Frank 2001; Nesse 2001). Batson, *et. al.* (1979) han explicado que los comportamientos altruistas de ayuda en situaciones de esta naturaleza están relacionados con la más que probable (y necesaria) reciprocidad intertemporal entre los miembros de un grupo estable que comparten una amenaza o un beneficio. Los experimentos de psicología social llevados a cabo por el equipo de Batson han demostrado que el altruismo se incrementa en situaciones de amenaza compartida, como ocurre con el entusiasmo militante de Lorenz. En estas circunstancias, el sacrificio altruista por el grupo resulta facilitado mediante procesos cognitivos y emocionales del tipo “conocimientos-y-sentimientos-compartidos”, que fortalecen la cohesión interna (*we feelings*). Asimismo, el denominado “efecto vecindario” explica también, en este caso apoyándose en la teoría de la selección inclusiva, la disposición al altruismo cuando existe una elevada probabilidad de que los individuos integrados en una

determinada población compartan una alta correlación genética y estén en situación de necesidad, o también cuando determinar el grado de parentesco no sea posible y la necesidad objetiva del destinatario del altruismo sea acuciante y evidente (Elster 2000[1979]: 240). En estos casos, dependiendo de la valoración del coste del altruismo respecto del beneficio supuesto que el mismo aporta al destinatario, un coeficiente de relación genética promedio, o incluso inferior, puede activar suficientemente la motivación para el comportamiento altruista.

La experimentación psicológica llevada a cabo acredita que la especie humana ha evolucionado hacia una psicología intersubjetiva con tasas de descuento interpersonal relativamente bajas en comparación con otras especies. Ello se debe a que el cerebro humano está más especializado que el de otros animales sociales en la cognición y motivación empáticas²⁵³. Además, tal y como se ha explicado en el apartado 3 de este capítulo, en la especie humana la modulación neurológica y psicológica de la empatía abarca un contorno social y relacional más amplio de congéneres que el puro nepotismo de la selección de parentesco. Esto prueba que más allá del altruismo parental existe un aplanamiento asintótico de la tasa de descuento interpersonal relativa a las múltiples afiliaciones grupales e identidades culturales del sujeto histórico.

Estas adaptaciones, que devienen en órdenes de cooperación y altruismo protoculturales, es más probable que aparezcan en aquéllas especies sociales dotadas de una mínima capacidad de replicar cultura. No obstante, aun en el caso de organismos con una gama elemental de innatismos, carentes de memoria a corto plazo y sin plasticidad fenotípica, es posible el surgimiento de una aparente actitud de preocupación por la comunidad sobre la base de un rudimentario mutualismo “de ensayo y error”. El modelo de interacción subyacente puede quedar correctamente ilustrado con el ejemplo del juego del destino común (*mutual fate control game*), que genéricamente se presenta como sigue en la figura 4.1.

²⁵³ Los experimentos clásicos de referencia son de Milgram (1974) y Batson (1991). La neurociencia cognitiva se ha adentrado con posterioridad en este campo, llegando a conclusiones similares [Lamm, *et. al.* (2007), Singer y Lamm (2009), Greene (2003), y Hauser (2006)].

Figura 4.1.- Matriz de pagos del juego del destino común

		Jugador 1		
		C	D	
C	+	+	-	Jugador 2
D	+	-	-	

La estructura del juego del destino común puede interpretarse como la expresión abstracta de incertidumbre social con equilibrios múltiples, donde los pagos quedan cuantitativamente indeterminados y se representan sólo mediante una sencilla escala binaria (+ -)²⁵⁴. En este juego puede evolucionar la cooperación si los jugadores pueden ser capaces de adaptar “a ciegas” sus conductas a la consecución de un destino común, por ejemplo a través de algún proceso inconsciente de aprendizaje adaptativo (genético o cultural) sometido a procesos selectivos. Aquí se entiende por cooperación el comportamiento (C,C) en el que ambos jugadores obtienen simultáneamente pagos positivos en la misma ronda, es decir, que ambos actúen como si les importaran favorablemente los buenos resultados del otro. Para ello no resulta necesario ni siquiera que los jugadores tengan noción alguna de lo que significa cooperar/defectar, o que los jugadores tengan conocimiento de que están inmersos en interacciones en las que existe una interdependencia *de facto*. Basta con que conozcan sus propias experiencias internas y las ordenen en función de preferencias estables.

Los experimentos de laboratorio realizados dentro del esquema de incentivos del juego del destino común han resultado esclarecedores acerca de la evolución de la cooperación. En efecto, como explica Colman (2005: 220-221), los experimentos realizados hace más de medio siglo por Sidowski (1957) y Sidowski, *et. al.* (1956) con jugadores humanos concluyen en un aprendizaje adaptativo de la cooperación²⁵⁵. Un primer experimento consiste en sentar a dos jugadores en habitaciones separadas, sin

²⁵⁴ Téngase en cuenta que la base de este juego es bastante cotidiana, dado que describe situaciones de racionalidad limitada e información incompleta. En el juego representado en la matriz de la figura 4.1. no hay estrategias dominantes, sino cuatro equilibrios posibles. Estamos pues en un contexto de indeterminación en el que puede surgir un punto focal por tanteo a través de algún proceso de aprendizaje filo u ontogenético. La selección (natural o cultural) puede estabilizar la cooperación a través de la implantación de instintos y/o la interiorización de normas culturales aprendidas.

²⁵⁵ Otros investigadores han repetido el experimento también con otros animales. Véase la literatura citada en Colman (2005: 221).

que tengan conocimiento el uno de la existencia del otro. A ambos se les instalan electrodos en su mano izquierda, y ambos tienen ante sí un par de botones para elegir estrategias que se adaptan al juego del destino común. Asimismo ambos disponen de un contador digital en el que aparecen los puntos obtenidos en cada jugada. Los jugadores reciben puntos cuando obtienen el pago (+) y una descarga eléctrica en su mano izquierda cuando obtienen el pago (-). Cuando pulsan el botón (C) el cojugador recibe puntos, y cuando pulsan el botón (D) el cojugador recibe una descarga eléctrica. El resultado del experimento indica que los jugadores tienden a coordinar impremeditadamente sus acciones en el sentido de la cooperación (el equilibrio C,C), incluso en este caso en el que ambos piensan que los pagos están relacionados con algún tipo de algoritmo misterioso que guarda relación con el azar o con su conducta pasada²⁵⁶. Después de alrededor de 200 repeticiones, la frecuencia con que los jugadores participantes en el experimento eligen cooperar (desconociendo el significado de esa elección) se sitúa en una horquilla de entre el 75%-80%. En otro experimento idéntico, salvo que ahora los jugadores fueron informados de que había otra persona en otro lugar cuya conducta determinaba tanto las descargas eléctricas como los puntos que obtendría (es decir, ambos son conscientes de la interdependencia estratégica), los resultados finales fueron muy similares a los del primer experimento. Esto implica que el conocimiento de la interdependencia mutua no es necesario ni influye de manera apreciable en la tasa de cooperación. Es decir, sea cual sea el mecanismo cognitivo-decisionario que explica la evolución de la cooperación, ésta puede producirse en condiciones más generales de lo que habitualmente se recoge en la literatura acerca de la reciprocidad.

Las condiciones en que supuestamente evolucionó la especie humana durante el Pleistoceno (interdependencia y poblaciones finitas) cumplen con las características propias de un marco de destino común de baja complejidad. Este marco conlleva un sistema de incentivos coherente con que los individuos realicen conductas altruistas, aunque carecieran de consciencia reflexiva o incluso su motivación fuera deliberadamente egoísta (Batson 1992: 72 y ss.)²⁵⁷. Un escenario experimental

²⁵⁶ Téngase en cuenta que ambos desconocen su interdependencia *de facto*.

²⁵⁷ Algunos trabajos de economía experimental proponen que la cooperación empírica que se aprecia en los juegos de dilemas sociales, ineficiente desde el punto de vista de que se aparta de la estrategia optimizadora, tiene su explicación evolutiva en los programas cognitivos adaptados para la

representativo de las condiciones sociales en que evolucionó el género *Homo* es el de la convivencia y los comportamientos de ayuda en la provisión de alimentos que se dan en las tribus actuales de cazadores-recolectores en Sudamérica. Basándose en las investigaciones bioantropológicas de estas tribus, Sarah Blaffer (2009: 14) explica que en todas las estudiadas se da un reparto comunal equitativo de la comida conseguida durante la jornada. Todos los individuos se alimentan aunque algunos no hayan tenido éxito en la caza o recolección. De acuerdo al estudio de los riesgos asociados a la consecución de alimentos en ese ambiente (algo que, obviamente, desconocen los indígenas), la conducta racional es compartir la comida con los demás miembros de la tribu. Aquellos individuos que fueran por libre y no compartieran la comida experimentarían un peligroso déficit calórico promedio en el 27% de los días de su vida. Además, al menos una vez cada 17 años los individuos que no comparten la comida se arriesgan a morir de hambre porque son incapaces de llegar siquiera al 50% de la ingesta calórica diaria necesaria para la supervivencia de un organismo humano (1.000 calorías). Cuando no se alcanza la ingesta mínima durante 21 días consecutivos se puede entrar en estado de coma por inanición. Sin embargo, cuando todos individuos de la tribu comparten la comida, el porcentaje de días en los que no se consigue ingerir ese mínimo de calorías se reduce desde el 27% a tan sólo el 3%, lo cual indica que el comportamiento de ayuda mutua es un rasgo adaptado de supervivencia²⁵⁸.

Parece claro que en circunstancias rutinarias el altruismo de las tribus puede ser explicado con la teoría convencional de la selección de parentesco y efecto vecindario, ya que es de suponer que en ellas los individuos compartían un elevado índice de semejanza genético. Sin embargo, conforme las tribus fueron incrementando su tamaño y produciéndose intercambios tanto de patrimonio genético como cultural, fueron apareciendo nuevos rasgos replicables compartidos que ampliaron el ámbito comunitario del destino común, como los que devienen de las relaciones reproductivas y afiliativas de tipo marital, vecinal y étnico-cultural (Wilson y Daly 1997: 260). El

cooperación en contextos de interdependencia y poblaciones finitas, o sea, un marco de destino común y superjuegos. Aun sin visibilidad o en ausencia de facultades cognitivas elevadas, en estas circunstancias un proceso de selección no muy exigente puede producir adaptaciones favorables a la cooperación en el juego del ultimátum anónimo de única jugada (Rand, *et. al.* 2013).

²⁵⁸ La única especie primate actual que se comporta con criterios de equidad intragrupal en la persecución y logro de objetivos conjuntos es la humana (Hamann, *et. al.* 2011). El experimento del juego del ultimátum revela que la equidad en el reparto puede evolucionar aun en ambientes de anonimato, siempre y cuando la población sea finita (Rand, *et. al.* 2013).

destino común de una población no se circunscribe exclusivamente a los objetivos comunes para la supervivencia a corto plazo de sus individuos. Los objetivos comunes giran en torno a la eficacia en la replicación de elementos que constituyen un cúmulo comunitario de replicadores tanto biológicos como culturales. Conforme las comunidades crecen en tamaño y complejidad se reduce el coeficiente de relación genética, pero puede aumentar la eficacia promedio de los individuos del grupo si resulta viable la acción colectiva y el funcionamiento del mercado. Por esta razón sus individuos están comparativamente mejor en términos de eficacia respecto a los de otras poblaciones de menor dimensión y complejidad. Además, el acto de altruismo marginal llevado a cabo es más eficiente en el primer caso que en el segundo, lo cual compensa el efecto incertidumbre asociado al ejercicio del altruismo dentro de una población de mayor heterogeneidad.

Estas fuerzas integradoras alimentan procesos de selección en varios niveles, que se explican por las teorías coevolutivas que consideran las consecuencias de la interacción de un *pool* de replicadores más numeroso y diversificado, y sus repercusiones en la eficacia expansiva de la población respecto a otras poblaciones²⁵⁹. La mayor riqueza y diversidad del patrimonio evolutivo que comparte una población supone una ventaja comparativa que puede explicar los procesos de expansión y extinción de las poblaciones. Éste pudo ser el caso de la interacción entre neandertales y cromañones durante 10.000 años, y la extinción de los primeros datada hace 30.000:

... (esta diferencia comparativa) habría permitido a los cromañones establecer vínculos entre su comunidad más allá de los biológicos, reconocerse no sólo por las relaciones de parentesco sino por las culturales. Y las relaciones culturales se extienden a lo largo de espacios más amplios que las biológicas. Un cromañón sería reconocido como tal por los símbolos de su cuerpo, y así posiblemente ayudado por otro cromañón desconocido de tierras lejanas. (Arsuaga 2001: 58).

Para Bowles y Gintis (2011: cap. 11) la clave de la organización social de relaciones altruistas está en los procesos de selección grupal reforzados por la coevolución genes-

²⁵⁹ El papel de la regulación cultural en la explicación del éxito de los grupos más grandes y complejos es crucial. Los grupos articulan la acción colectiva mediante estructuras organizativas (dotadas de reglas, incentivos, sistemas de supervisión, etc.) que seleccionan conductas en orden a los resultados e intereses comunes. Esto requiere la capacidad de discriminar las conductas de optimización individual. La cuestión clave aquí es que los individuos que cooperan dentro de estructuras con un diseño adecuado obtienen mejores resultados individuales que si van por libre. Véase Simon (2001).

cultura expresada a través de las emociones sociales. Esto conecta con la formación de identidades fenotípicas no basadas en la consanguinidad, sino en un patrimonio replicable más amplio (de carácter comunitario) que posibilita la evolución de grupos de altruistas integrados y fortalecidos por la inclusión enculturada de *outsiders* (Richerson y Boyd 2005; Schmid 2005b; Efferson, *et. al.* 2008).

La clave a considerar entonces se plantea como una cuestión de replicadores mixtos (genotípicos y fenotípicos), que en sí mismos constituyen un impremeditado conjunto de objetivos compartidos determinantes de un destino común. Se trata de la perspectiva evolutiva poblacional, que igualmente permite explicar la evolución del altruismo sin necesidad de consanguinidad. La perspectiva poblacional aporta eficacia comparativa al individuo a través de relaciones de cooperación y altruismo, pero no obstante la evolución de estos comportamientos requiere adaptaciones específicas que eliminen o reduzcan los efectos negativos en la eficacia individual derivados de la exposición a comportamientos no cooperativos. El estudio de este fenómeno parte de la existencia de equilibrios entre estrategias de comportamiento no individuales, sino poblacionales. Su análisis se aborda a continuación.

4.6.- La perspectiva poblacional

Tanto entre especies como dentro de una misma especie existen relaciones de cooperación y conflicto. De acuerdo a una determinada concepción del darwinismo algunos enfoques metodológicos asumen que las fuerzas de selección operan en la línea exclusiva de una deriva completamente homogeneizadora. Esta visión sostiene que sólo existe un estereotipo representativo de los individuos de la especie, porque las demás alternativas ya han sido (teóricamente) superadas por las diversas formas de competición en procesos de selección intraespecífica. En economía, por ejemplo, sólo existe un estereotipo de agente representativo, el *homo oeconomicus*²⁶⁰. En la que hemos llamado paradoja de Darwin (capítulo 3), los agentes racionales habrían llevado a la extinción a aquellos otros irracionales o cuasi-racionales. Una expresión empírica que

²⁶⁰ Los enfoques dentro del individualismo metodológico se apoyan en la existencia de un único agente representativo: el actor racional. Sin embargo, los estudios empíricos sobre esta temática revelan una manifiesta diversidad dentro de la especie y la cultura. Una manifestación principal de esta diversidad empírica es la que emana de los rasgos de identidad egoísmo *versus* altruismo (López y Valdaliso 1999: 33 y ss.)

confirma esta tesis es que personas con un perfil psicológico sociopático, que manifiestan no sentir remordimientos de conciencia por practicar defección en dilemas sociales y morales, y que revelan frialdad emocional, agresividad y una baja aversión al riesgo, tienden a ocupar puestos relevantes en la jerarquía social y obtener resultados en los negocios superiores a la media (Koenigs, *et. al.* 2007; Shiv, *et. al.* 2005). Siendo esto así podría pensarse que este perfil psicológico, y su arquitectura de preferencias, deberían haberse impuesto en la evolución por su mayor eficacia relativa. Sin embargo esta conclusión es teórica y empíricamente falsa (Thaler 2000: 136).

La observación de la naturaleza nos lleva a descubrir que los individuos de una especie no son un *continuum* absolutamente homogéneo, sino que coexisten diferencias internas significativas (grupos de rasgos) en aspectos relevantes de la morfofisiología y comportamiento. Tipologías sutilmente diferentes o muy diferentes de comportamiento en la interacción intraespecífica son formas de adaptación coherentes con una supervivencia dentro de equilibrios poblacionales. En lo que corresponde a la coexistencia en la especie humana de individuos con diferentes tendencias hacia el egoísmo y el altruismo psicológicos, o a la defección y la cooperación respectivamente, numerosos estudios demuestran la simultaneidad de tendencias congénitas polarizadas hacia expresiones más prosociales o antisociales del carácter (Kendler, *et. al.* 2007; Meyer-Lindenberg, *et. al.* 2006; Kreek, *et. al.* 2005; Hare 2006; Mednick, *et. al.* 1984). Asimismo se ha documentado la existencia de culturas y grupos etnográficos con tendencias relacionales más cooperativas y sociedades más igualitarias y cohesionadas que otras (Malinowsky 1926; Boehm 1993; Konner 1982). En un contexto evolutivo la heterogeneidad de motivaciones e inclinaciones de comportamiento puede ser explicada por diferencias culturales y educativas entre las personas, así como por disposiciones genéticas (Wade 2015).

Dentro de otro campo de estudio, la investigación neurocientífica ha encontrado pruebas neurofisiológicas y neuroanatómicas explicativas de la disposición congénita hacia comportamientos honestos, generosos y altruistas, o bien deshonestos, egoístas y aprovechados (Morishima, *et. al.* 2012; Greene y Paxton 2009; Rand, *et. al.* 2012). Asimismo ocurre con la tendencia a seguir en la interacción social motivaciones de tipo impulsivo-emocional o bien de carácter calculador – utilitario (Koenigs, *et. al.* 2007;

Greene 2004). Por otro lado, la economía y la psicología experimentales ha presentado resultados que avalan una estructura poblacional humana estable, integrada por tres patrones de comportamiento social: cooperadores condicionales, altruistas y defectores (Grujic, *et. al.* 2010; Kurzban y Houser 2005; Fischbacher, *et. al.* 2001). En esta línea, los estudios de evolución simulada realizados por Lomborg (1996) permiten demostrar la posibilidad teórica de este tipo de equilibrios poblacionales en relación con la cooperación. En su investigación Lomborg muestra la estabilidad evolutiva de un equilibrio poblacional semejante al hallado por la economía experimental, en este caso compuesto de los que él llama cooperadores incondicionales, cooperadores precavidos y no cooperadores. Conclusiones parecidas son las obtenidas en el experimento de Back y Flache (2006), en el que se aprecia cómo la interacción de patrones conductuales alternativos respecto a la cooperación genera espontáneamente una estructura ordenada de relaciones dentro de la población. Esta dinámica posibilita que el rasgo cooperación sobreviva dentro de una estructura poblacional polimórfica.

Reconocida la diversidad de caracteres en lo relativo a la cooperación, y dado que la cooperación es un rasgo en aparente desventaja evolutiva respecto a la defección, es preciso analizar: (a) cómo confluyen las poblaciones hacia este tipo de equilibrios evolutivos estables; y (b) de qué adaptaciones adicionales disponen los individuos para poder señalar y detectar eficazmente su naturaleza y evitar, de este modo, ser explotados por los no cooperadores.

4.6.1.- El equilibrio evolutivo entre egoístas y altruistas

Las investigaciones realizadas por Tomasello, *et. al.* (2012) y Tomasello y Vaish (2013) permiten demostrar que, además de la capacidad para un lenguaje complejo y un gran cerebro dotado de capacidad para generar autoconciencia, la especie humana se diferencia de las demás de primates sociales en que es la única adaptada biológicamente para participar con éxito en acciones colectivas que requieren cooperación. La cooperación llevada a cabo con éxito se erige sobre la base de metas e intenciones comunes, y planes de acción socialmente coordinados.

Los experimentos realizados con interacciones grupales de individuos humanos demuestran que los grupos más homogéneos respecto a la actitud cooperativa de sus integrantes generan tasas de cooperación más elevadas en la acción colectiva, así como una ganancia individual superior a la que se produce en el caso de mezclar disposiciones conductuales heterogéneas²⁶¹. El talante amistoso y capaz de perdonar agravios pasajeros consigue mejores resultados de eficacia que las actitudes vengativas, competitivas y pendencieras (Fudenberg, *et. al.* 2012; Dreber, *et. al.* 2008). De esto se infiere la base evolutiva para explicar la preferencia por la cooperación y, dentro de ella, por cooperar selectivamente con otros cooperadores. Sin embargo, no se excluye por ello la coexistencia de rasgos antitéticos, como la cooperación pacífica y la defeción agresiva. De hecho, la heterogeneidad de rasgos dentro de una población evoluciona hacia equilibrios polimórficos en que los pagos derivados de las distintas tipologías de comportamiento tienden a igualarse (Maynard Smith 1982).

El comportamiento prosocial humano se manifiesta como si los individuos actuaran a favor del grupo (e incluso de la especie) y no exclusivamente de sí mismos. Esto exige admitir la existencia de procesos de optimización selectiva en los que los beneficiarios trascienden al individuo-organismo²⁶². Tal hipótesis tiene una larga trayectoria, y se halla presente en la obra de Adam Smith en el s. XVIII²⁶³. Permite explicar, por ejemplo, un comportamiento extremo de auto-inmolación aparentemente disfuncional como el heroísmo cívico, militar y filantrópico, o las conductas terroristas

²⁶¹ En el experimento de Kurzban y Houser (2005: 1.807) los grupos de cooperadores en los que no se incluye ningún *free-rider* generan una ganancia individual de en torno al 40% superior que la obtenida si se introduce una mínima heterogeneidad. La formación de grupos homogéneos (de cooperadores o no) mejora el funcionamiento de la interacción y los resultados de la misma y además, en el caso de cooperadores de alto compromiso, permite la estabilidad evolutiva de este rasgo (Gächter y Thöni 2005; Burlando y Guala 2005).

²⁶² La disposición a los equilibrios poblacionales depende de lo que algunos autores han denominado ecología del comportamiento, y que aquí llamamos racionalidad ecológica o perspectiva poblacional. Según Boyd y Silk (2004: 552) la ecología del comportamiento humano consiste en «la teoría del actor racional llevada un paso más allá. Predice que los intereses de los individuos se definen en términos evolutivos: los humanos quieren maximizar su eficacia biológica». Esta “pretensión” conlleva que no siempre se actúe optimizando individualmente en los dilemas sociales. Se trata de casos como que sea preferible, desde la perspectiva de la población de rasgos homogéneos, optar por una optimización compartida con otros (compartir la comida con los miembros del grupo), o simplemente cedida a otros (alimentar primero a los niños cuando no hay comida para todos).

²⁶³ Concretamente lo que planteó Smith fue la siguiente conjetura:

En todo el universo vemos cómo los medios se ajustan con esmerado artificio a los fines que están destinados a producir; y en el mecanismo de una planta o un cuerpo animal admiramos cómo cada cosa es diseñada para alcanzar los dos mayores propósitos de la naturaleza, *el mantenimiento del individuo y la propagación de la especie*. (Smith 1997[1759]: II, ii, 3, p. 187; cursivas añadidas).

de los suicidas atacantes, también llamados “mártires mortíferos” (Tobeña 2005). Los kamikazes japoneses o los yihaidistas radicales, entrenados para convertirse en mártires letales, inmolan sus vidas en acciones mortales de ataque que obviamente no les benefician a ellos, sino a amplios agregados poblacionales en favor de los cuales ellos se sacrifican en nombre de una construcción cultural como la patria o el pueblo de Dios. Este comportamiento carece de sentido evolutivo, salvo que la progenie del suicida altruista reciba de su sociedad favores y honores extraordinarios, o que con él se produzca un reequilibrio poblacional entre dos especies, linajes, etnias o grupos culturales con identidades y objetivos excluyentes²⁶⁴. Un ejemplo en animales no humanos es la estrategia altruista programada genéticamente en las mariposas nocturnas de la familia *Saturniidae*. Estos insectos exhiben una coloración aposemática tóxica que llama la atención de sus depredadores. Las mariposas señalizan, a costa de la propia vida, que el conjunto de la especie es peligrosa para el depredador (Dugatkin 2007: 183-184).

Igualmente, la perspectiva poblacional admite la existencia de emociones y sentimientos adaptados para motivar comportamientos intrafamiliares también aparentemente disfuncionales, como el infanticidio, pero que analizados desde la perspectiva poblacional cobran sentido evolutivo:

... la selección natural favorece los mecanismos fisiológicos que acaban con el embarazo cuando las posibilidades que tiene el feto de sobrevivir son bajas. La misma lógica adaptativa se aplica a los niños una vez han nacido. El razonamiento evolutivo sugiere que los padres deben estar predispuestos a cuidar solamente aquella descendencia que tiene probabilidades de sobrevivir y reproducirse satisfactoriamente. (Boyd y Silk 2004: 571).

Comportamientos de esta naturaleza invitan a introducir en el modelo de racionalidad la perspectiva poblacional, en la que se admite que existen motivaciones, preferencias, objetivos y perspectivas cognitivas adaptadas para la optimización de agregados poblacionales (Ginges y Atran 2009; Colman, *et. al.* 2008a; Fritz y Singer 2008; Schmid 2007; Bowles 2006).

²⁶⁴ Un modelo evolutivo de referencia, explicativo de este tipo de comportamientos de rivalidad malintencionada o maliciosos (*spiteful behaviors*) es el de Hamilton (1970). En Price, *et. al.* (2002) se contrasta esta teoría, aplicándola a la evolución de sentimientos de castigo moral como adaptación equilibradora de la eficacia relativa consecuencia de la interacción entre cooperadores y *free-riders*. Un ejemplo de este patrón de conducta es la aplicación de represalias costosas a los *free-riders* en interacciones de última jugada.

Desde la evidencia del comportamiento altruista, Dawkins no encuentra ninguna razón convincente para explicar estos comportamientos de apariencia anormal sin asumir el gen como unidad analítica básica de replicación. A diferencia del individualismo metodológico, la propia teoría evolutiva de la selección de parentesco explica la conducta (y su motivación) desde una perspectiva supraindividual:

Sin mirar la vida con los ojos del gen no hay razón particular alguna por la que un organismo deba “cuidar” su éxito reproductor y el de sus parientes en lugar de, por ejemplo, atender a su propia longevidad. (Dawkins 1993[1976]: 301-302).

La propagación de “los parientes”, en función de lo que se mide el éxito reproductor de un organismo, nos emplaza de lleno en la perspectiva de la población de individuos que comparten replicadores (genéticos, y también culturales). Así lo propone Gintis explicando la evolución del comportamiento altruista:

... no existe conflicto alguno entre el cálculo de la eficacia individual y los cálculos a nivel grupal o poblacional, ya que admitiendo que ese cálculo existe, debemos tener en cuenta el hecho de que las pérdidas de eficacia que se derivan de los actos altruistas individuales pueden ser compensadas a un nivel social más elevado, de modo que los genes causales del altruismo no están en contradicción con el nivel poblacional. (Gintis, *et. al.* 2008: 248)²⁶⁵.

Las relaciones de cooperación y conflicto dentro de un ecosistema obedecen a la interacción y estructura de los diferentes grupos homogéneos de replicadores. La perspectiva poblacional ubica los organismos en una carrera armamentista para extraer eficacia a través de relaciones de cooperación y altruismo, así como de estrategias ofensivas y defensivas frente a otros organismos (Levine 1998; Price, *et. al.* 2002). Toda carrera armamentista evolutiva es costosa para sus contendientes, y a la larga aparecen fuerzas que conducen a detenerla propiciando un equilibrio evolutivo estable con resultado polimórfico (Maynard Smith y Price 1973). El polimorfismo estructural tiende a la estabilidad hasta que surja alguna innovación que rompa dicho equilibrio y reactive la carrera armamentista. Como apunta Gifford (1999: 134), desde una perspectiva propia de la teoría de la elección pública la presión selectiva favorable a cerebros más grandes y complejos puede explicarse como resultado de una búsqueda de rentas evolutiva: el desarrollo de instituciones que limiten los costes de dicha competición

²⁶⁵ Lógicamente, esto no es compatible con el individualismo metodológico, para el que los individuos son la única unidad de análisis, y generan equilibrios sociales buscando optimizar una eficacia no inclusiva.

vendría a ser una búsqueda de rentas biológicas en la arena cultural.

Existen abundantes ejemplos en la literatura especializada que evidencian esta teoría. Así, tal y como explica Robert Frank, las fuerzas competitivas que impulsan los atributos diferenciales (“rasgos de poder”) de los machos en la selección sexual, pueden acabar detrayendo eficacia a los individuos más conspicuos. Esto ocurre con los ejemplares mejor pertrechados para la competición intraespecífica de leones marinos, ciervos, pavos reales, etc., que resultan más vulnerables al ataque de depredadores. El reequilibrio podría haberse conseguido con mayor rapidez si los machos actuaran como si regularan racionalmente el poderío de sus atributos individuales. Es decir,

... si todos los machos pudiesen reducir (sus atributos de poder) a la mitad, todos y cada uno de ellos saldrían beneficiados. Los ganadores de las batallas seguirían siendo los mismos, pero todos serían menos vulnerables a los depredadores [...] en la lucha, lo que importa es el tamaño relativo (...) si todos fuesen más pequeños, el resultado de cada combate sería el mismo y todos (...) disfrutarían de mayor protección frente a los depredadores. (Frank 2008b: 25 y 136-137; paréntesis añadido).

En lo concerniente a soluciones culturales humanas igualmente encaminadas a limitar las consecuencias disfuncionales derivadas de la lucha intraespecífica por la eficacia diferencial, los sistemas humanos refuerzan la cooperación mediante regulaciones institucionales orientadas a reducir los efectos externos que conducen a tragedias de acción colectiva. En los procesos competitivos en que los premios individuales dependen en gran medida del rendimiento relativo (y no absoluto), lo más normal es convenir reglas formales o informales que refuerzan la cooperación. Frank cita algunas de ellas, como por ejemplo la edad a la que los niños comienzan el colegio, la obligatoriedad de los uniformes escolares, la normativa sobre seguridad laboral, o, incluso, la prohibición de la poligamia:

... intente encontrar un ejemplo de actividad organizada que premie el rendimiento relativo de los competidores, pero que no haya intentado evitar las inversiones en mejoras de rendimiento que se anulan mutuamente. Si hay alguna, yo no la conozco (...) Observe las normas que establecen los diferentes grupos humanos e intente averiguar cuál puede ser su finalidad. (Frank 2008b: 274).

Estas soluciones cooperativas (biológicas y culturales) a la problemática asociada a las carreras armamentistas son Pareto-eficientes, en su caso, desde la perspectiva

individual, grupal y de especie (Maynard Smith y Price 1973). Dasgupta (2000: 71) considera las normas morales formas de “fricción” que restringen las posibilidades del comportamiento. Esta regulación es un resultado evolutivo complejo (no deliberado) que posibilita que los comportamientos sean predecibles y, de este modo, acotar la gama de equilibrios posibles. Según este planteamiento, el fundamento de toda sociedad está en tales restricciones a la indeterminación.

Otro ejemplo ilustrativo que puede mencionarse es el de los comportamientos de crianza cooperativa de la progenie. La monogamia se da en numerosas especies de aves y algunas de mamíferos. Cuando no se dispone de adaptaciones específicas para el reconocimiento de la progenie, como ocurre con la mayoría de los animales, en las especies diploides la monogamia conlleva costes directos y de oportunidad medidos en eficacia. Si los emparejamientos son entre individuos monógamos, sin embargo, los beneficios reproductivos superan esos costes y por ello la monogamia puede ser una ventajosa adaptación evolutiva en especies como la humana (Daly 1978)²⁶⁶. En estos casos queda claro que existe un objetivo de destino común (la progenie que porta el ADN de la pareja), sólo y en la medida que el comportamiento promiscuo esté autolimitado, lo cual es coherente, en el caso humano, con la existencia adicional de restricciones culturales a la promiscuidad. Esto es coherente asimismo con una perspectiva poblacional que es compatible con la maximización de la eficacia individual. Los promiscuos (halcones) se comportan como si sólo se interesaran por su propia replicación, aprovechándose de la fidelidad de los monógamos (palomas). Sin embargo, cuando abundan los promiscuos sus probables emparejamientos constriñen su eficacia hasta que vuelven a abundar monógamos gracias al superior éxito reproductivo de su comportamiento honesto.

El modelo de restricciones a la libertad de acción individual se aprecia en el comportamiento reproductivo de otras especies, incluso vegetales, que igualmente parecen comportarse como si los individuos asumieran voluntariamente compromisos

²⁶⁶ Este enfoque teórico parte de las investigaciones de Trivers (1972) sobre la inversión parental. Gifford (1999: 137-138) utiliza la teoría del éxito reproductivo para explicar la evolución de la exclusividad sexual en especies de mamíferos carnívoros, y su relación con la participación de los machos en la alimentación de las crías. Otro estudio reciente conecta el éxito reproductivo de las especies de primates monógamos no humanos con la implicación del macho en la defensa de sus crías ante el ataque infanticida de otros machos (Opie, *et. al.* 2013).

autoimpuestos²⁶⁷. Asimismo se refleja en la ejecución de los patrones humanos de dominación y liderazgo²⁶⁸. ¿Qué ventaja adaptativa obtiene el organismo que se comporta como si autorrestringiera su comportamiento a favor de un agregado intraespecífico más amplio? Admitiendo que los individuos que cooperan tienen mucho que perder frente a un posible comportamiento oportunista que tuviera éxito entonces, *ceteris paribus*, para que pueda sobrevivir la cooperación es necesario que no cooperar esté asociado a una distribución de probabilidad que conduzca a situaciones de una eficacia reproductiva inferior a la resultante de cooperar. De este modo, el individuo que decidiera ir por libre para coger con el paso cambiado a los demás, e intentara aprovechar las ventajas individuales del oportunismo, saldría perdiendo puesto que, buscando su beneficio individual, se enfrentará a una elevada probabilidad de fracaso que afecta tanto a él mismo como a su descendencia. En situaciones en que los costes individuales del egoísmo pueden ser muy altos, proliferará la población de altruistas hasta alcanzar un equilibrio entre ambos rasgos. Los individuos que cooperan siguen una racionalidad ecológica como si se percatasen de que cooperando consiguen mejores resultados de eficacia poblacional, coherentes con la expectativa de su propia eficacia. Siendo esto así, los individuos semejantes “optan” por un comportamiento prudente y conformista que genera un orden social de cooperación²⁶⁹.

El fenómeno se representa matemáticamente con una función de la eficacia que integra la estructura de incentivos correspondiente a un comportamiento orientado a la selección del organismo a través del oportunismo, o bien a la selección del grupo de organismos (*pool* de rasgos) mediante la cooperación. Puede ilustrarse con una función lineal del tipo de la utilizada en los experimentos sobre financiación de bienes públicos:

²⁶⁷ Este sería los casos de los bambúes (*Phyllostachys bambusoides*) y las llamadas cigarras periódicas, o langostas (*Magicicada Septendecim*) que expone Gould (2010[1977]: cap. 11).

²⁶⁸ Puede consultarse el modelo de Noë y Rebello (1994) aplicado al equilibrio entre los estilos honestos y deshonestos en la forma de dirigir empresas. Las redes de relaciones humanas se hacen y se desacen, como es el caso de grupos informales y organizaciones formales, en función de la dinámica experimentada en la invasión de defectores y su capacidad de replicarse. El modelo desarrollado por Cavaliere, *et. al.* (2012) demuestra que los ciclos vitales de las organizaciones pueden explicarse por la dinámica de equilibrios entre cooperadores y defectores, el ocaso de aquéllas dominadas por los segundos, y la ruptura del contrato social por parte de los cooperadores cuando son incapaces de excluir o evitar a los defectores.

²⁶⁹ El comportamiento prudente y conformista se producirá siempre que la interacción con el medio determine que $p(dc > 0) > p(dnc > 0)$, siendo p la probabilidad de que el rasgo “cooperar” (c) o “no cooperar” (nc) produzcan una descendencia superior a cero.

$$F_i = x_i + a \cdot \sum_{(j: 1, n)} c_j ,$$

donde F es la eficacia obtenida por el individuo i de una población n , derivada de invertir una cantidad de recursos m en bienes privados (x) y/o en bienes de disfrute conjunto (c), de tal manera que para cada individuo participante $m_j = x_j + c_j$. El parámetro a está situado en una escala tal que $0 \leq a \leq 1$, y representa un *coeficiente de cooperatividad*, o medida de la inclinación que tiene cada individuo a comportarse de manera progrupal. Lógicamente el egoísta realiza una aportación progrupal $c = 0$ como estrategia dominante, puesto que para él $a = 0$. Sin embargo, el cooperador hace una aportación $c > 0$, ya que su estrategia dominante es cooperar en función del valor que él tenga de a . Para los individuos con un valor de $a = 1$ la estrategia dominante consiste en $m = c$, dado que la estructura de pagos del juego está determinada de tal manera que $(n \cdot \bar{a}) > 1$, siendo \bar{a} la cooperatividad promedio, para que la cooperación sea la conducta Pareto-eficiente desde la perspectiva grupal.

En una estructura de incentivos como la especificada, el coeficiente de cooperatividad de los no cooperadores siempre es próximo o igual a 0, mientras que el correspondiente a los cooperadores siempre es próximo o igual a 1. Para simplificar el análisis asumimos solamente los casos límite, 0 y 1. De este modo, se plantea que las interacciones pueden ser eficientemente aprovechadas con o sin cooperación ($a = 1$ y $a = 0$ respectivamente). En las interacciones entre ambos tipos de individuos los defectores persiguen la optimalidad individual propia de ganadores en juegos de suma cero, mientras que los cooperadores actúan como si buscaran una optimalidad paretiana poblacional de juegos de suma positiva. Desde el punto de vista evolutivo la selección natural puede favorecer los dos tipos de conducta, tanto la característica de los depredadores y oportunistas, como la propia de cooperadores. Sin embargo, obsérvese que si tratamos de procesos evolutivos comparados, en interacciones típicas de acción colectiva los organismos capacitados para cooperar *exclusivamente* con otros cooperadores obtendrían mejores resultados que los defectores en la interacción social. Esto es altamente relevante para explicar la evolución de la cooperación, puesto que si el comportamiento predatorio u oportunista fuese siempre más eficiente para el individuo que la cooperación, los procesos de selección natural hubieran conducido a un mundo habitado sólo por defectores, es decir, un mundo de unimórfico de no

cooperación. Ese hipotético escenario no es la naturaleza real, y menos aún el mundo humano que se caracteriza, en mayor o menor grado, por el predominio de la cooperación y el rechazo cultural de los comportamientos predatorios y oportunistas recalcitrantes. Hasta tal punto es así que los patrones de conducta sociopática no sólo son excepcionales (por considerarse patológicos y anormales) en las relaciones intraespecíficas, sino que además en las sociedades humanas son castigados, desprestigiados y evitados, limitándose de este modo su replicación (Richerson y Boyd 2001: 219).

Es muy difícil sobrevivir como paloma en un ambiente de halcones, sobre todo si las palomas se muestran incapaces de diferenciar a otra paloma de un halcón. Consecuentemente, la estrategia predatoria devendría expansiva hasta invadir a toda la población, un resultado que no es coherente con los modelos evolutivos de interacción estratégica entre depredadores y presas (halcones y palomas). La conclusión consiste en que si hay palomas (es decir, si existe empíricamente la cooperación) es porque ha podido evolucionar, y si esto es así, se debe a que las palomas normalmente se relacionan con palomas, pueden reconocer a los halcones y evitar ser depredadas. Si los psicópatas obtienen mejores resultados en interacciones arriesgadas en las que sentir culpa o compasión es un obstáculo, ¿por qué no es psicópata toda la población? Tal vez porque la característica ausencia psicopática de escrúpulos y habilidades sociales incapacita para comprometerse en procesos grupales de cooperación. O porque sus rasgos de personalidad espantan a los cooperadores, que se han adaptado para evitarlos, aislarlos, desprestigiarlos y advertir de su condición a los demás cooperadores. El cooperador puede disfrutar de una lucrativa reputación a largo plazo entre sus afines, y ello aunque se trate de un compromiso que a corto plazo lo aboque a situaciones desfavorables frente a perfiles sociopáticos de conducta (Back y Fache 2006).

Los modelos descriptivos de este tipo de procesos evolutivos se inspiran en los trabajos iniciales de Maynard Smith y Price (1973) y Maynard Smith (1974, y 1976), así como en Frank (1988: cap. 3) y en Dugatkin y Wilson (1991). Aquí nos apoyamos en la versión ofrecida por Sober (1996: 223 y ss.), que presentamos resumidamente²⁷⁰.

²⁷⁰ Otras versiones explicativas de la misma teoría pueden consultarse en Bergstrom (2002), Frank (1994: 239-245), y Dawkins (1993[1976]: cap. 5).

Supongamos que existen dos estrategias de comportamiento dentro de un ecosistema poblacional que conforman dos grupos de rasgos: los explotadores (E), que extraen recursos de los demás y que nunca comparten sus recursos; y los cooperadores (C), que siempre comparten sus recursos. La distribución de resultados que propician las diferentes combinaciones de individuos cuando interactúan son los que figuran en la siguiente matriz de pagos:

Cuando interactúa con

		Explotador	Cooperador
		Explotador	-25
Cooperador	0	15	

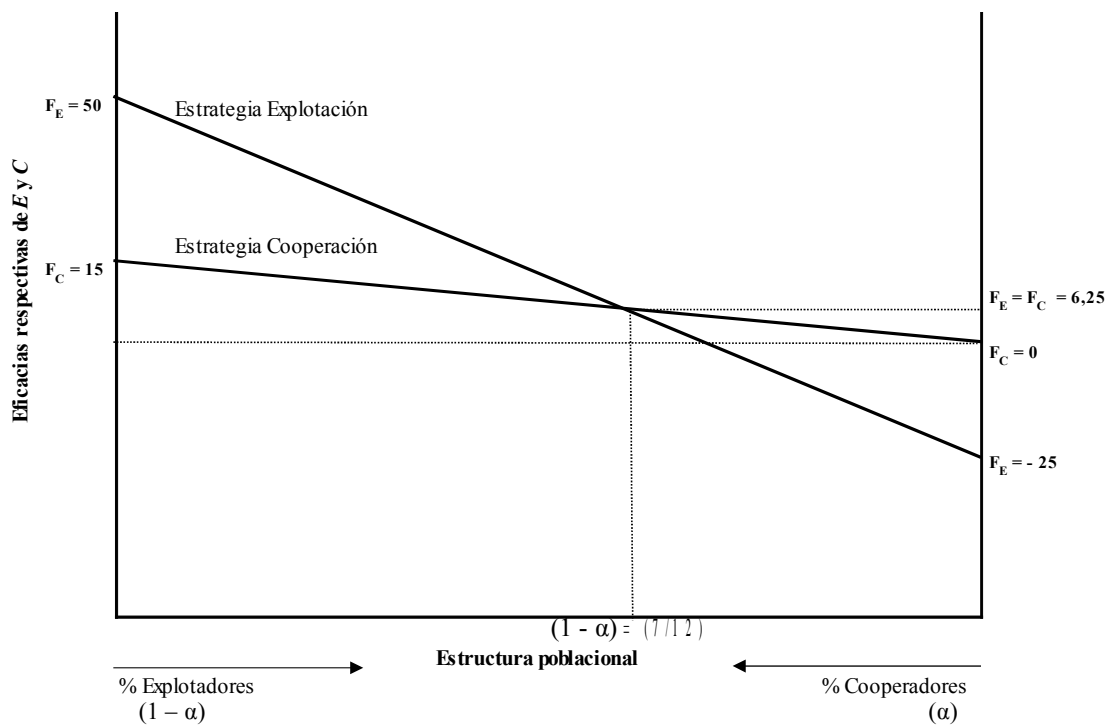
Resultados del...

De este modo, cuando un *E* interactúa con otro *E* el resultado promedio es una pérdida de eficacia valorada en -25, dado que el cruce de estrategias de explotación desemboca en promedio en un conflicto de suma negativa. Por el contrario, cuando *E* se cruza con *C* el resultado es una ganancia de 50, ya que *E* explota fácilmente a *C*, que suponemos nada obtiene de esta interacción (resultado igual a 0). El cruce de *C* con otro *C* supone que ninguno pugna por conseguir más que el otro, y comparten equitativamente el resultado conjunto de su interacción, que suponemos un valor de 30 distribuido a partes iguales. La expresión matemática de esta matriz de pagos se plasma en dos ecuaciones de resultados que dependen de la probabilidad de interactuar con cada uno de los dos rasgos. Si suponemos que la proporción de *C* en el ecosistema es α , y que por lo tanto la proporción de *E* en ese mismo ecosistema es $(1 - \alpha)$, y siendo los emparejamientos aleatorios, la ecuación de eficacia esperada de cada uno de los dos rasgos es:

$$F_E = 50 \cdot \alpha - 25 \cdot (1 - \alpha) ;$$

$$F_C = 15 \cdot \alpha$$

Figura 4.2.- Representación del equilibrio evolutivo estable de explotadores y cooperadores



La representación gráfica de ambas ecuaciones (figura 4.2.) permite visualizar cómo en un ambiente dominado por cooperadores la presencia de un solo explotador marginal le aporta una excepcional ventaja adaptativa (50) frente a la condición de cooperador (15). Debido a ello, en sucesivas generaciones los explotadores irán reproduciéndose diferencialmente respecto a los cooperadores y su frecuencia en la población será creciente (con una eficacia decreciente). Por el contrario, si el cien por cien de la población estuviera constituido por explotadores, el resultado de la interacción entre ellos les restaría eficacia (-25). Un solo mutante cooperador obtendría un resultado marginal mejor (0) que permitiría replicarse con mayor éxito relativo, impulsando el crecimiento de la proporción de cooperadores en el conjunto de la población (con eficacia creciente). En el punto de corte de ambas rectas -con una proporción de $(1 - \alpha) = 7/12$ de explotadores, y $\alpha = 5/12$ de cooperadores- se produce un *equilibrio evolutivo estable* (o *estado evolutivo estable*) en el que ambos grupos de población obtienen una eficacia esperada idéntica, valorada en 6,25. Si no se producen cambios en los parámetros de las ecuaciones, estos resultados tienden a permanecer estables, aunque a corto plazo puedan darse alteraciones circunstanciales de la estructura poblacional. Cualquier alteración es necesariamente pasajera, dado que no se trata de una estrategia

evolutivamente estable²⁷¹. En cuanto se da una leve alteración entran en funcionamiento fuerzas selectivas que conducen a un nuevo equilibrio evolutivo dentro del dimorfismo estructural.

La perspectiva poblacional demuestra que la estrategia cooperación puede evolucionar aunque en la interacción de los cooperadores individuales con la estrategia defección salgan comparativamente desfavorecidos en términos de eficacia. Para que la cooperación tenga presencia estable en un ecosistema social es suficiente con que aparezca un solo mutante cooperador capaz de replicarse a sí mismo. De este modo, la cooperación puede sobrevivir a través de un mecanismo indirecto de eficacia inclusiva no radicado en el enclave familiar. Si la cooperación con otros cooperadores es probable y provechosa, y existen adaptaciones que les permiten señalar la cooperación y discriminar a los defectores, entonces el que de vez en cuando el cooperador individual se cruce con algún defector no impide que el rasgo pueda evolucionar dentro de la población. La eficacia derivada de cooperar con otros cooperadores compensa con creces el que se cuele algún defector ocasional no excesivamente dañino (Fletcher y Doebely 2009).

Consecuentemente, las fuerzas de selección que posibilitan el equilibrio evolutivo de estrategias en una población, son en sí mismas un mecanismo de selección individual coherente con la selección del grupo de vehículos portadores de ese mismo rasgo. En grupos con más cooperadores, especialmente si los grupos compiten entre ellos con diferentes estructuras internas de cooperación/defección, les va evolutivamente mejor a los cooperadores. Los individuos que cooperan pueden manifestar un interés especial por la supervivencia de sus afines. Puede tratarse perfectamente de la optimización inconsciente de un *pool* genético o cultural del que participa el individuo, y para cuya replicación es preciso desarrollar adaptaciones prosociales:

²⁷¹ Téngase en cuenta que aunque la interacción de ambas estrategias conduce a la aparición de un equilibrio evolutivo estable, sin embargo esta condición no implica que alguna de las dos estrategias sea una estrategia evolutivamente estable (EEE). El que en un momento dado se diera una estructura poblacional dominada al cien por cien por cualquiera de las dos estrategias no garantiza que no pueda ser invadida por la estrategia alternativa. Consecuentemente, no se trata de una EEE. El equilibrio evolutivo estable es una propiedad de una población integrada por más de un rasgo estratégico; no así la EEE, que se trata de un atributo individual que determina el comportamiento. Véase Sober (1996: 224 y 227).

... se puede decir que alguien coopera en el sentido de que manifiesta una actitud o voluntad cooperativa hacia alguna meta o actividad colectivamente promovida, y hacia los que participan de ella. (Tuomela 2000: 5).

Cuando la competencia social de los individuos resulta determinante para el logro de un objetivo interdependiente (como es el caso de la supervivencia del rasgo “capacidad para la coordinación en pos de una meta común”), parece lógico que evolucione la especial valoración de aquellos individuos de los que se tienen indicios verosímiles de ser honrados compañeros (Tomasello y Vaish 2013: 239). Esto confiere una ventaja selectiva a los buenos compañeros respecto a los deshonestos, y dado que ello beneficia a los individuos dotados del fenotipo honestidad, resulta verosímil la hipótesis de que evolucionen adaptaciones para señalar el rasgo y reconocerlo en los afines, así como para detectar y experimentar aversión ante la percepción del fenotipo defección.

4.6.2.- Selección discriminatoria de fenotipos

La mayoría de modelos evolutivos comparten una premisa común sin la cual el altruismo no pueden evolucionar en su interacción con la estrategia egoísta: ha de existir algún modo de criba para que la interacción no sea cognitivamente aleatoria; o sea, se ha de poder reconocer y supervisar a los individuos explotadores a fin de tener capacidad de evitarlos (Ostrom 2000: 145). Esto equivale a que en ambientes sociales inciertos los costes de transacción no superen el coste del riesgo asociado a cooperar a ciegas. La selección a favor de la afinidad es un factor de evolución de los rasgos afines, siempre que la interacción con los afines (*affinity assortment*) aporte ventajas de supervivencia diferenciales. Como exponen Hoffman, *et. al.* (1998: 349), una teoría de la mente aplicada al comportamiento cooperativo necesita incluir una explicación de algún tipo de adaptación que permita el reconocimiento diferenciado de “amigos” y “enemigos”, así como la detección de las intenciones ajenas. El acceso a información sobre la predisposición, motivación e intencionalidad de otros individuos o grupos de individuos es una capacidad evolucionada que permite seleccionar aquéllos dispuestos a cooperar, que están comprometidos con el comportamiento cooperativo. Los altruistas genuinos pueden de este modo cooperar con sus afines, pero además los procesos de selección inclusiva del altruismo posibilitan el ejercicio de la reciprocidad positiva, un refuerzo adicional del rasgo altruismo que, en sí mismo, es un factor de evolución.

Las capacidades evolucionadas para la discriminación de fenotipos son necesarias para que las adaptaciones prosociales, eficaces para la supervivencia en grupo, sean viables eludiendo la explotación por parte de otros individuos. La respuesta adaptativa a comportamientos no cooperativos forma parte de la carrera armamentista dentro de un ecosistema social que tiende al equilibrio evolutivo. Se trata de adaptaciones que cualifican un modelo concreto de psicología animal:

El deseo de tomar ventaja a través de un desapercibido oportunismo puede ser una tentación básica de la naturaleza humana, y el deseo de cooperar preferiblemente con otros cooperadores -y de confirmar su disposición cooperativa a través de información sobre su reputación o supervisión directa- puede estar del mismo modo profundamente enraizado en la evolucionada psicología humana. (Price 2006: 214).

Consecuentemente,

Determinados dispositivos neurocomputacionales antiexplotación son características de diseño centrales e indispensables de cualquier estrategia cooperativa, especialmente si la estrategia pretende emerger con éxito en la población y permanecer estable. (Tooby, *et. al.* 2006: 107).

Esta hipótesis ha de ser contrastada mediante su rastreo en la psicología humana actual así como, en un nivel de evolución comparada, en otros primates sociales genéticamente cercanos. En este sentido, algunos de los principales hallazgos de la economía experimental aplicada a la primatología, muestran que en especies evolutivamente cercanas al género *Homo* existen tendencias conductuales a la reciprocidad, la supervisión mutua y una preferencia por la división equitativa de los resultados de la acción colectiva. Estas adaptaciones posibilitan a los individuos extraer eficacia de la interacción social sin desviarse hacia la pérdida o ganancia extraordinaria de ventajas por la vía del engaño y la explotación.

Los estudios de psicología comparada revelan que al menos dos especies de primates sociales no humanos de la actualidad, que son los chimpancés y los monos capuchinos, disponen de adaptaciones de esta naturaleza²⁷². En la observación de estos animales en estado natural se aprecia que los resultados de la acción conjunta (p.ej. la caza) tienden a ser repartidos equitativamente entre los participantes y no así a los que no participan o lo hacen remoloneando. Los que pretenden aprovecharse de los demás

²⁷² Para una ampliación de lo expuesto a continuación véanse De Waal (2005), De Waal y Berger (2000), y Brosnan y De Waal (2003).

son rechazados y amonestaos, y se evita volver a actuar con ellos en interacciones futuras. En experimentos de laboratorio se interpreta la arquitectura emocional de estos animales. En efecto, ambas especies, igual que la humana, están capacitadas para cooperar y profundizar la cooperación, con aquellos que juegan limpio en la acción conjunta de manera reiterada. Asimismo, se interrumpen las relaciones o aparecen reacciones de rabia, indignación y desprecio hacia los que no se muestran equitativos o son sorprendidos realizando conductas desleales. Los monos que cooperan en el logro de un objetivo común (p.ej. atraer con unas cuerdas en paralelo alimentos, bien dentro de una única bandeja o en compartimentos separados) tienden a repartir equitativamente lo conseguido, y cuando aparece una recompensa desequilibrada a la acción conjunta (uvas para uno, el alimento predilecto de los capuchinos, piña para otro) los monos muestran airadas reacciones de indignación y rabieta, ya sea contra el compañero o el experimentador, en función de su capacidad de atribuir (a propósito, con bastante buen criterio) intenciones y causalidad al resultado no equitativo. Hasta tal punto puede llegar la reacción emocional de la aversión a los comportamientos no equitativos, que los monos suelen rechazar seguir participando en el experimento, y no aceptan la comida que se les proporciona, si aprecian que se está recompensando desigualmente una participación individual que no es desigual.

Diera la impresión de que existe un trinquete evolutivo dentro del mecanismo de destino común: cuando una especie se beneficia de las ventajas de ser cooperativa, surgen refuerzos adaptativos de la perspectiva poblacional, como puede ser el que podemos decir instinto del trabajo bien hecho, que incluye la exigencia de que los demás cumplan equitativamente su parte²⁷³. Esto se explica porque de no reforzarse la equidad, la honestidad, la reciprocidad y la gratitud, por ejemplo, la cooperación tiende a degenerar, y si esto ocurre, a la larga (y a veces a muy corto plazo), acaba teniendo efectos sistémicos que reducen la eficacia potencial de la especie en su conjunto. Se trata pues de una arquitectura psicológica adaptada para reforzar la cooperación a través del programa de cooperación con (y premio a) los cooperadores.

En los estudios llevados a cabo con humanos se ha encontrado una evidente

²⁷³ Agradezco la apostilla sobre el instinto del trabajo bien hecho al director de esta investigación.

correlación entre la cantidad y calidad de interacción comunicativa y los indicadores de cooperación. Los jugadores que intervienen en experimentos de laboratorio tienden generalmente a cooperar en mayor medida, y de manera sostenida, cuando pueden verse o identificarse unos a otros, aun sin comunicarse de ningún otro modo (Wichman 1970; Burnham 2003; Bohnet y Frey 1999a). Asimismo, cooperan con bastante más frecuencia e intensidad cuando pueden comunicarse verbalmente o por escrito, ya sea o no directamente cara a cara (Ostrom, *et. al.* 1992; Davis y Holt 1993; Sally 1997; Ledyard 1995; Frey y Bohnet 1996; Frank 2006; Bochet, *et. al.* 2006, y Brosig, *et. al.* 2003), y aunque la comunicación se realice mediante escuetos mensajes a través de terminales de ordenador y sin contacto visual (Kimbrough 2008; Sell y Wilson 1991). Por otra parte, se ha detectado que las personas cooperan en mayor medida cuando saben que están interaccionando con otras personas y no con un programa informático (Rilling, *et. al.* 2002 y 2004a). Frank, *et. al.* (1993b) constatan, apoyándose en la metodología del dilema del prisionero con jugada única que, cuando se permite conversar presencialmente durante un breve tiempo a los participantes, se produce una elevada tasa de acierto en la predicción sobre quiénes no cooperarán. Estos contactos comunicativos previos resultan eficaces en aras de incrementar la cooperación, puesto que, *cæteris paribus*, se eleva en la primera jugada desde el 20 – 30% hasta el 70 – 80% cuando se permite a los jugadores sostener conversaciones de 30 minutos antes de empezar a interaccionar (Frank 2006).

Aunque no es posible contar con una información cien por cien fidedigna acerca de las intenciones ajenas, y eso es preciso reconocerlo, sin embargo en la vida empírica los individuos humanos suelen obtener unos resultados satisfactorios cuando se fían de su intuición en las interacciones sociales²⁷⁴. Existen pruebas de que el cerebro dispone de dispositivos especializados (emplazados en el córtex prefrontal intermedio y lateral) en la decodificación de la información que permite inferir el conocimiento de las intenciones ajenas (Haynes, *et. al.* 2007). La inteligencia intuitiva procesa mucha más información -a nivel inconsciente- que el intelecto racional (Gladwell 2005), y aunque se admite una carrera armamentista entre las estrategias de honestidad y deshonestidad

²⁷⁴ Indica Simon (1983: 36) al respecto que: «gran parte del pensamiento humano y del éxito que tienen los seres humanos al obtener decisiones correctas se debe al hecho de que cuentan con buena intuición o buen juicio».

comunicativa, vamos a explicar cómo en ésta existe una dinámica propia de la tendencia al equilibrio evolutivo que se basa en no concebir la actividad mental ajena como una caja negra²⁷⁵. ¿Es posible reconocer experimentalmente un acto altruista en sus coordenadas de motivación no-utilitaria? La señalización de una identidad honestamente cooperadora, así como la capacidad de identificar a los defectores y falsos cooperadores son adaptaciones necesarias para que el equilibrio poblacional no se rompa. Por ello, la teoría respalda que tales adaptaciones son verosímiles:

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 (...) Existe amplia evidencia para soportar el argumento de que los humanos responden a actos altruistas de acuerdo a su percepción de los motivos del altruista. Se tiende a responder más altruistamente cuando se percibe al otro actuando de manera “genuina”, esto es, realizando voluntariamente un acto cooperativo como un fin en sí mismo, sin esperar ganancia. (Trivers 1978[1971]: 216 y 218; cursivas en el original).

Pasemos a analizar los principales resultados de los estudios realizados al respecto.

4.6.2.1.- Estrategias de señalización

Las estrategias de señalización de una identidad cooperativa pueden ser congénitas o aprendidas, y sirven al individuo que las ejecuta para expresar el fenotipo de sus replicadores de forma identificable por sus afines, además de mostrar el compromiso con determinada conducta u objetivo de forma inequívoca, de manera que influya en el comportamiento de otros individuos, tanto cooperadores como defectores.

La primera propiedad de la señalización, revelar replicadores, aporta ventajas adaptativas desde la perspectiva del ejercicio directo de la eficacia inclusiva, como en el caso de la certeza de estar cooperando con un pariente consanguíneo. Dado que la teoría evolutiva parte de la base de que todo organismo que tenga la certeza de que buenas copias de sus replicadores (genéticos y culturales) “viajan” en otro organismo será capaz de sacrificarse por él si la población de réplicas gana eficacia con ello (Dawkins 1993[1976]: 115), entonces un rasgo de comportamiento, como el altruismo mismo, saldrá reforzado en la evolución si quienes lo replican son capaces de señalar que lo

²⁷⁵ Una reflexión sobre la problemática de la orientación doctrinal conductista en psicología, y sus implicaciones y problemática asociada a la caracterización de la naturaleza humana, puede encontrarse en Andreoni, *et. al.* (2008).

poseen, así como de reconocer en otros organismos esta misma posesión²⁷⁶. El *pool* de replicadores homogéneos, de este modo, se estará ayudando a sí mismo a través de uno de los rasgos replicables que integran el *pool*, que es el comportamiento altruista²⁷⁷.

Para que la operación replicadora sea eficiente se requiere que exista algún fenotipo perceptible que señalice la posesión de cualquier rasgo común, y específicamente el del altruismo²⁷⁸. Una forma eficaz de señalar el carácter altruista consiste, obviamente, en realizar actos altruistas (Dawkins 1993[1976]: 117). Si el organismo que posee cualquier rasgo obtiene una ventaja evolutiva derivada de señalarlo fidedignamente, el organismo altruista se adaptará en orden a señalar sus cualidades para maximizar su eficacia. Cuanto mayor dificultad exista en la falsificación de las señales por parte de competidores de calidad altruista inferior, más éxito evolutivo tendrá el altruista en su eficacia (Frank, *et. al.* 1993). Asimismo, cuanto mayor facilidad exista en la confusión, más ventaja adaptativa obtendrán aquellos que lleven su conducta a un altruismo extremo. Señales de gran generosidad, incluso con desconocidos, como la contribución a la provisión de bienes públicos altamente costosa, son difíciles de falsificar por defraudadores, dado que se trata de una señal verificada en sí misma que equivale a una externalidad positiva no internalizable. Los resultados de los experimentos de laboratorio y trabajos de campo realizados por Gneezy, *et. al.* (2012) se muestran concluyentes al respecto: las conductas prosociales que no son costosas para el individuo señalizador carecen de suficiente credibilidad en los destinatarios de la señal. Por ello, la señalización a través de contribuciones onerosas a bienes públicos, u otros actos altruistas en el sentido fuerte, se muestran como una estrategia adaptativa eficaz (Gil-White y Richerson 2003: 833).

En este aspecto, Vernon Smith (1998: 8) sostiene que: «El comportamiento no cooperativo consiste en la ausencia de reciprocidad positiva, no en la práctica de la

²⁷⁶ El *pool* de replicadores comunes, de este modo, puede evolucionar aunque los organismos se degraden o desaparezcan como unidades ontológicas:

... una señal de altruismo real permite la selección de grupo. Permitir la selección de un grupo similar supone la perpetuación de individuos similares, así como su aislamiento de los que son diferentes. (López y Valdaliso 1999: 36).

²⁷⁷ Si el altruismo puede replicarse con mayor efectividad mediante la transmisión cultural, como así es, entonces la propagación a través del propio altruismo es un fin intrínseco.

²⁷⁸ Esta es la llamada teoría de la señalización a través de “etiquetas fenotípicas”, popularizada por Richard Dawkins con su famosa metáfora señalizadora consistente en una llamativa barba verde (Dawkins 1993[1976]: 117).

reciprocidad negativa», o sea, no cooperar equivale a comportarse sin equidad. Esto supone que quien no coopera es el que actúa aprovechándose de los demás cooperadores sin practicar la reciprocidad positiva con ellos, una forma de explotación que incluye el acto de extraer recursos ajenos de manera violenta y sin contraprestación, como en la conducta predatoria. Dada la naturaleza de la motivación específica del depredador y el parásito, estas tipologías de comportamiento encuentran dificultades para inhibir la gratificación psicológica asociada a una oportunidad de aprovecharse de la vulnerabilidad ajena (Greene y Paxton 2009)²⁷⁹. Cualquier señal altruista, como propagar gratuitamente un recurso a la comunidad, puede recibir reciprocidad positiva o no. Si ocurre lo primero se trata de un indicio de que se está en un ambiente de cooperadores. Si la señal altruista recibe como respuesta saqueo u oportunismo, entonces se infiere que se está en un ambiente en el que hay defectores. De ello, si en el dilema del prisionero de jugada única el cooperador que lanza una primera señal de cooperación recibe reciprocidad positiva y no existe posibilidad de error, entonces se infiere que con alta probabilidad el cojugador es otro cooperador.

La estrategia de señalización pública de una identidad altruista orientada a la replicación ha sido explicada por Simon (1983: 60-61), basándose en los estudios de E. O. Wilson sobre una determinada especie de insectos acuáticos. A veces algunos de estos individuos mutan, y para propagar la mutación buscan afines señalizando su disposición reproductiva mediante una estrategia altruista. Concretamente, secretan una sustancia que disuelta en el medio acuático contribuye a reducir la acidez del agua, y esto beneficia a los insectos semejantes que viven en ese ecosistema (una charca). La consecuencia directa del comportamiento señalizador altruista es que durante el corto período de apareamiento del que disponen estos insectos se hallan en un medio que les provee de recursos y bienestar, y además la señal atrae y diferencia, es decir, identifica a los afines, con el objeto de reproducirse con ellos y propagar sus replicadores (incluida la mutación) con mayor rapidez que el resto de pobladores del ecosistema. Si la estrategia tiene éxito, de modo que los mutantes lleguen a reproducirse a una tasa superior que los demás subgrupos, entonces el rasgo tenderá a incrementar su frecuencia en la población hasta el límite que marque el nuevo equilibrio evolutivo.

²⁷⁹ No aprovechar la vulnerabilidad ajena para tomar ventaja es un indicio de comportamiento moral, y de capacidades para llevar a cabo acciones de cooperación.

A parte de estos testimonios acerca de cómo funciona la naturaleza en los procesos de evolución, y aunque hemos de tener presente que la especie humana no es ajena a ellos, sin embargo no debemos pasar por alto que los homínidos se valen especialmente de la cultura para señalizar. Para ello se acometen patrones de relación con procesos dirigidos y pautados con un sentido idéntico al de los insectos acuáticos descritos. En el caso de la especie *Homo sapiens* las innovaciones, ya sean de comportamiento, de conocimiento o materiales, son la señal que, transmitida y recibida por aquéllos a quienes pueden beneficiar, permite conformar grupos con una tasa de crecimiento superior a la de otros grupos. Esto ocurre sobre la base de la rápida identificación de individuos que son semejantes desde una perspectiva fenotípica, que puede ser orgánica o psicológica, es decir, identitaria. Dicho esto, añaden López y Valdaliso:

El altruismo real de carácter genético está presente en los humanos, pero es el de contenido cultural el que más hemos desarrollado frente al resto de homínidos. El mayor acto de altruismo real (cultural) que podemos hacer los seres humanos es poner a disposición de la comunidad una creación (tecnológica, institucional, científica o artística) (...) en realidad estamos mandando una señal para identificarnos y crecer rápidamente como grupo, asegurándonos una “descendencia” similar. (López y Valdaliso 1999: 37).

Una vez constituido el grupo, el siguiente paso es favorecer su supervivencia y replicación mediante ulteriores innovaciones que otorguen a los que comparten esa identidad biológica y/o cultural ventajas frente a los *outsiders*. Tales ventajas se concretan en capacidades diferenciales y benéficas para los *insiders*, y que no sean fácilmente copiables por los no afines.

La posibilidad de hacer falsas copias de fenotipos señalizadores de una identidad, con el objeto de obtener ventaja de ello, supone un problema para los señalizadores de una genuina identidad altruista. Para que una señal de calidad sea fiable es preciso que no pueda ser fácilmente utilizada por impostores²⁸⁰. Dentro de la carrera armamentista entre señalizadores honestos y deshonestos puede resaltarse que los primates son especialmente diestros (y sofisticados) para mentir, simular y

²⁸⁰ En el plano puramente biológico, un ejemplo de señal creíble sobre la calidad de la salud y de la fertilidad, por ejemplo, son los múltiples atributos corporales que se incluyen en el atractivo físico de las personas (Luxen y Van de Vijver 2006), algunos de los cuales pueden, incluso, suponer una desventaja frente a agentes infecciosos o depredadores (Hamilton y Zuk 1982). Un marcador veraz en la revelación de una procedencia cultural puede ser el acento al hablar (Cohen 2012).

tergiversar estratégicamente, como es el caso paradigmático sublimado de los caracteres psicopáticos humanos, u otros casos que con el oportuno entrenamiento llegan a ser altamente persuasivos en estas habilidades (p.ej. los actores profesionales). Sin embargo, Paul Ekman (1985 y 2003) argumenta que la evolución de la carrera armamentista ha consolidado expresiones no falsificables de estados emocionales e intencionales, las cuales no son reproducibles fácilmente con intenciones falsificadoras. Además, si la evolución ha equipado con la capacidad de simular y tergiversar a mutantes falsificadores, igualmente lo ha hecho con los señalizadores fidedignos, dotándolos de capacidades adaptadas para significarse a sí mismos, así como detectar y neutralizar fraudes y defraudadores (Ekman 2005; Ectoff y Ekman 2000; Cosmides y Tooby 1992 y 2005)²⁸¹.

Este aspecto de la problemática asociada a la veracidad de las señales enlaza con la otra cualidad de la señalización mencionada. Dicha cualidad se orienta a mostrar preferencias y compromisos inequívocos, de modo que la información transmitida influya en el comportamiento de otros afectando sus expectativas (Shelling 1978c: 229). La especie humana es estratégica por antonomasia, y buena parte de los comportamientos sociales ritualizados, es decir, de la cultura, carece de sentido si no se interpreta como señalización estratégica. Muchos de estos comportamientos con frecuencia son costosos para quien los practica, aunque, admitiendo que no siempre son conscientes, no por ello son irracionales o incluso provechosos para el organismo que los realiza. En esta línea se encuadran los comportamientos consistentes en manifestaciones somáticas y conductuales de las emociones, tanto más significativos y fiables de una intencionalidad no susceptible de falseamiento como difíciles sean de controlar de forma voluntaria (Frank 1988).

Otras particularidades del proceso señalizador de carácter estratégico, adaptadas para evitar el problema de la veracidad cooperativa, tienen que ver con el contenido y la especificidad del mensaje transmitido con la señal. Así, *cæteris paribus* una señal tiene una probabilidad tanto mayor de ser veraz en la señalización a través de la interacción

²⁸¹ La sinergología es la ciencia que se encarga de conocer lo que comunica el cuerpo de manera inconsciente y paralela a la expresión verbal y a la simulación gestual: «cuando el ser humano disfraza la verdad, no se siente nunca plenamente a gusto (...) Las manos, el cuerpo, no se han sometido al examen de la lógica; sin embargo, son el subtítulo de las palabras» (Turchet 2010: 27).

social si: (a) la señal es cognoscible y/o aprovechable sólo por parte de cooperadores; (b) el contenido explícito de la señal requiere una fuerte inversión previa de parte del señalizador (una inversión previa de alta especificidad si no se utiliza para señalar el compromiso inequívoco con el comportamiento altruista); y (c) en ocasiones la emisión de la señal está acompañada de consecuencias personales irreversibles. Cuando se dan especificidades e irreversibilidades, si la señalización no tiene éxito entonces el valor de la inversión realizada en la señal puede perderse, de lo cual se deduce la existencia de una actitud de compromiso fuerte con una determinada identidad altruista²⁸².

El objeto de muchas donaciones altruistas es aflorar e identificar rápidamente a los afines para cooperar con ellos y evitar la tragedia comunal. En el plano empírico una señal inequívoca de portar auténticamente el rasgo altruismo puede ser ejercer un altruismo pertinaz, intensivo y recurrente, costoso, y sin posibilidad de un beneficio previsible derivado de esta conducta a corto o a largo plazo. Este vendría a ser el caso de estrategias de señalización del tipo “quemar las naves”, como las de los misioneros, cooperantes sin retorno, ciertas uniones civiles, o personas que donan todo su patrimonio a una determinada causa. La señalización a través de un recurso sólo valorable por los afines es la donación altruista de algo que permite detectar a los semejantes porque sólo ellos (los que son altruistas como el donante) lo aprecian. Este podría ser el caso de los sistemas voluntarios y altruistas de donación de sangre, órganos y médula ósea (López y Valdalisó 1999: 36)²⁸³. Aquí la donación sólo puede interesar a destinatarios con una fracción compartida de ADN y, dada esta consideración, se trata de individuos filogenéticamente semejantes, y por ello con una probabilidad superior de ser igualmente altruistas. Los que no son semejantes (y de ello improbablemente altruistas) no tienen incentivo a aprovecharse de este altruismo, luego por esta vía los defectores y oportunistas no pueden extraer eficacia detrayéndola del altruista. Por otra parte, si donante y receptor no son anónimos, el estímulo a cooperaciones recíprocas y gratificaciones ulteriores es evidente. La tercera expresión de veracidad de la señal, la

²⁸² Rodney Stark denomina a estas cuestiones de incentivos “el problema de la credibilidad”, y lo asocia a la ética del sacrificio y la estigmatización dentro de procesos culturales de iniciación. Véase Stark (2001: cap. 8). Trivers (1972) incluye este análisis de la señalización dentro de los procesos de selección sexual entendida como inversión parental en recursos específicos.

²⁸³ Valgan también los ejemplos de la entrega gratuita de un objeto sólo apreciado por los seguidores de un determinado credo religioso o político, o la aceptación de una invitación voluntaria a un acto en el que los asistentes realizarán actos de altruismo.

de la especificidad de la inversión necesaria para poder emitirla, puede consistir en la adquisición de habilidades costosas, como las de una extensa carrera académico-profesional, que estén completamente especializadas en la realización de acciones altruistas. Este sería el caso de los largos procesos de iniciación y capacitación para la asunción de responsabilidades en ciertos ámbitos de la vida académica, familiar, religiosa, política, militar o profesional, estrechamente relacionados con el ejercicio del altruismo.

Pero el éxito del proceso de señalización se enmarca en una relación entre emisor y receptor. Una señal es eficaz si se emite y llega en condiciones adecuadas a otro individuo que, lógicamente, ha de estar capacitado para comprender la señal e identificar la naturaleza de su emisor. Ello incumbe a la capacitación para descifrar las auténticas intenciones de la señalización y la genuina naturaleza del organismo señalizador. Pasemos pues a reflexionar acerca de las capacidades adaptadas para la identificación discriminativa.

4.6.2.2.- Reconocimiento de cooperadores y explotadores

Que los individuos *necesiten* cooperar para el logro de *sus* objetivos no es un argumento necesario ni suficiente para que se produzca la cooperación. Un ejemplo es el apareamiento y crianza de la progenie. La mayoría de los animales se aparean “como si” lo hicieran libre y voluntariamente, y muchos de ellos crían y sacan adelante a su progenie (especialmente las hembras), lo cual no quita para que también existan “violaciones” y comportamientos oportunistas, como los del cuco, que pone sus huevos en nido ajeno para que sean inadvertidamente empollados por otras aves huésped. Como adaptación reductora de la indeterminación, la naturaleza ha seleccionado una extraordinaria diversidad de expresiones de señalización de intenciones en el reino animal, que revelan tanto la disposición al ataque como un estado favorable al encuentro cooperativo, y a veces delatan incluso una utilización oportunista de falsas señales (Espmark, *et. al.* 2000)

Generalmente los animales muestran comportamientos de apariencia cooperativa en el cortejo, apareamiento y crianza de la prole. Esa cooperación resulta ventajosa para

el logro del objetivo de reproducirse. Para ello se valen de procesos ritualizados (innatos y aprendidos) que facilitan la detección y reconocimiento de cooperadores, de modo que el apareamiento eluda ser una interacción a ciegas susceptible de ser parasitada por un defector, un equilibrio no cooperativo al estilo del dilema del prisionero, o un equilibrio indeterminado en el caso de equilibrios múltiples. Las conductas de señalización y reconocimiento de la intencionalidad cooperativa tienen alto nivel de éxito en la detección de emparejamientos cooperativos viables y, una vez detectado el cojugador con el que es viable la cooperación, no se producen defecciones ni indeterminaciones en la selección del mejor equilibrio desde la perspectiva de la pareja. Eberhard Weismann describe estos procesos de señalización indicando que:

Durante la búsqueda de su pareja los animales se encuentran con muchas otras especies animales parecidas o cercanamente emparentadas con la suya. El observador no puede por menos que quedar admirado y fascinado ante la increíble seguridad (...) con que no sólo se encuentran los animales de la misma especie sino con que, incluso ente sus congéneres reconocen a la pareja distinguiéndola del rival. Esta seguridad en el reconocimiento es una condición esencial para que el apareamiento tenga lugar sólo entre animales de la misma especie y no se produzcan mezclas ni cruces entre las especies animales más distintas. (Weismann 1994: 51).

Los procesos sociales de reproducción biológica no son los únicos casos de acción colectiva en que los organismos reconocen, seleccionan y se vinculan con cooperadores en orden a conseguir objetivos conjuntos. Si la acción conjunta aporta mayor eficacia que la acción individual, existe una base explicativa de la cooperación y de otras adaptaciones especializadas para que esta sea posible. Así, muchos animales manifiestan capacidad para evaluar la calidad en las relaciones intraespecíficas (Nöe y Hammerstein 1994). Los primates sociales, por ejemplo, disponen de capacidades cognitivas adaptadas para recoger información socialmente relevante, lo cual les posibilita discriminar en sus relaciones. Asimismo revelan preferencias por la calidad, tanto en la elección de parejas de apareamiento como en la selección de los socios más leales:

... reconocen a otros individuos; identifican a sus parientes; computan el valor de los recursos y servicios; llevan un registro del resultado de las interacciones pasadas con el resto de miembros del grupo; realizan inferencias transitivas; discriminan entre cooperadores y defectores, y evalúan las cualidades de sus posibles rivales, parejas y aliados. (Silk 2007: 1.348).

Una vez reconocida y demostrada la calidad cooperativa de las relaciones empíricas entre individuos, resulta verosímil que dado que la calidad diferencial es un bien por naturaleza escaso, otras adaptaciones refuercen este ventajoso capital social acumulado de manera que sea posible consolidar alianzas estables y vinculaciones específicas²⁸⁴.

Tales capacidades evolucionadas están asociadas igualmente a la importancia del capital social en el rango ocupado y el éxito reproductivo. Las jerarquías de los primates no humanos son fundamentalmente de dominación, y la eficacia de los machos alfa está en función de la fuerza bruta y la astucia (Henrich y Gil-White 2001: 167). Las relaciones de cooperación suelen ser exclusivamente con parientes consanguíneos (hermanos de madre, y en mucho menor grado hermanos de padre). En el género *Homo*, que conserva parte de este patrimonio cognitivo social primate, sin embargo se aprecia además una jerarquía de tipo tutelar, basada en la calidad altruista del líder (*ibid.*). En la especie humana las capacidades cognitivas que posibilitan estas diferencias radican en un superior tamaño y complejidad del área cortical -concretamente el córtex medial prefrontal- donde se gestiona neurológicamente la faceta cognitiva de las relaciones de cooperación. La selección previa, coordinación, vigilancia y evaluación final de las complejas relaciones sociales de cooperación humanas exigieron adaptaciones más evolucionadas que las existentes en otros primates. Estas capacidades tienen que ver con la innata capacidad humana de realizar continuamente complejas inferencias sobre pensamientos e intenciones de otros individuos, comprendiendo y ubicándose desde la perspectiva y estado informativo de éstos (McCabe, *et. al.* 2000, y 2001). Los humanos disponen de una teoría de la mente que posibilita una interacción social mucho más compleja que la del resto de primates sociales, y que permite descifrar con suficiente precisión si el comportamiento ha sido premeditado o accidental²⁸⁵.

²⁸⁴ Véase Ostrom (2003). Por otra parte, McCabe, *et. al.* (1996) han demostrado experimentalmente, y Pacheco, *et. al.* (2008) teóricamente, que el éxito en las cooperaciones previas incrementa los resultados positivos de la futura cooperación a través del apego vinculatorio. La vinculación entre individuos transforma la matriz de pagos de la interacción, de modo que la cooperación con afectos vinculatorios se materializa en una actitud de preferencia por el individuo vinculado (actitud de confianza hacia el *old helping partner*). Esto a su vez estabiliza la relación cooperativa e incrementa paulatinamente los resultados de la misma.

²⁸⁵ Existen evidencias de que sólo la defección premeditada se procesa emocionalmente a través de circuitos diferenciados a los de la no cooperación accidental. El primer tipo es considerado una muestra de actitud, mientras que el segundo no resulta definitorio de un comportamiento tramposo u oportunista, y por tanto no merecedor de censura. Véase Price (2006: 208), y Greene, *et. al.* (2009).

Dado que el capital social de las relaciones entre individuos aporta eficacia diferencial, de ello el valor explicativo de adaptaciones cognitivas y emocionales especializadas en: (a) identificar y clasificar la calidad cooperativa de los individuos; (b) discriminar en el trato con ellos; y (c) apartarse/represaliar, o vincularse con, en función de la calidad demostrada. El modelo propuesto por Axelrod y Hamilton (1981) explica la evolución de la cooperación sobre la base de capacidades avanzadas de los individuos para detectar señales identificativas de una disposición cooperativa. Los cooperadores que tienen esta capacidad incrementan su eficacia, los que no, la reducen. Este planteamiento ha dado pie a modelos más refinados y una prolija literatura sobre verificación de cooperadores (*partner verificación*) mediante capacidades evolucionadas para el reconocimiento de la identidad cooperativa, así como el aprendizaje adaptativo y cambio de socios en función de los resultados obtenidos de la cooperación²⁸⁶.

El interrogante de fondo se refiere a que puesto que el cerebro funciona por defecto para la sinceridad, y existen procesos neurológicos diferenciados cuando se realizan comportamientos deshonestos intencionalmente (Gazzaniga 2006: 119), ¿es posible acceder a esos procesos mentales desde otra mente? ¿Expresa el organismo alguna señal perceptible al exterior de lo que está ocurriendo en su interior? En otro orden de cosas, si el comportamiento altruista está causado por la actividad en los circuitos paralímbicos, especializados en la producción de hormonas y otras sustancias neuromoduladoras de la confianza, vínculo, afecto amoroso y cooperación, cuya presencia en el organismo se acompaña de sensaciones de armonía, bienestar y euforia (Moll, *et. al.* 2006; Kosfeld, *et. al.* 2005; Rilling, *et. al.* 2002; Carmichael, *et. al.* 1994), ¿resulta posible detectar señales externas de esa actividad orgánica y así identificar a los verdaderos altruistas? Esto es altamente relevante para contrastar la hipótesis del reconocimiento de la veracidad cooperativa.

Existe actualmente cuantiosa evidencia experimental que demuestra la

²⁸⁶ Los trabajos se refieren al comportamiento animal desde la perspectiva sociobiológica, y algunos de ellos están específicamente orientados al comportamiento humano desde la antropología y la etología humana. Dentro de los modelos de detección y verificación pueden citarse los trabajos de Axelrod y Dion (1988), Boyd (1992), Dugatkin, *et. al.* (1992). Entre los modelos de evaluación y aprendizaje cabe citar a Eshel y Cavalli-Sforza (1982), Dugatkin y Wilson (1991), Friedman y Hammerstein (1991), Enquist y Leimar (1993), entre otros.

existencia de capacidades de reconocimiento, evaluación y selección de cooperadores. Estas capacidades son de tipo filogenético, y se manifiestan en la ontogénesis tanto a nivel conductual como en el plano psicológico y neurológico. Así, los niños pequeños de tres años demuestran experimentalmente ser capaces de diferenciar a los colaboradores de los explotadores en las acciones de provisión conjunta de recursos. Su comportamiento respectivo se corresponde con los patrones lógicos de un modelo innato de teoría de la mente que propicia la selección inclusiva del rasgo cooperación (Melis, *et. al.* 2013)²⁸⁷.

En esta línea cognoscitiva, el organismo humano se muestra filogenéticamente adaptado para reconocer señales de afinidad y familiaridad, o sea, para personalizar las relaciones. El reconocimiento (cognitivo y emocional) de rostros y objetos familiares se registra en áreas específicas de los lóbulos temporales y del córtex prefrontal. Asimismo, la piel emite una determinada tipología de señales electrogalvánicas ante la presencia de conocidos, amigos y parientes²⁸⁸. Igualmente, registra procesos neurológicos específicos cuando se miente o se da respuestas engañosas, cuando se reconoce algo verdadero y cuando se detectan mentiras y actitudes fraudulentas en la interacción social (Gazzaniga 2006: 119; Hoffman, *et. al.* 1998).

Una tesis central en la psicología evolutiva es que además de la disposición de competencias sociales como la cooperación y la reciprocidad, la especie humana manifiesta una alta especialización filogenética en competencias cognitivas para la detección de tramposos y explotadores, independientemente de la cultura a la que pertenezcan los individuos (Sugiyama, *et. al.* 2002). Los experimentos de Brown y Moore (2000) muestran además un éxito contundente en el reconocimiento e identificación de altruistas. Los individuos con una disposición más proclive a la confianza y el trato amistoso pueden ser más vulnerables a las artimañas de los defectores, salvo que manifiesten, como es el caso, una habilidad más acertada que la de

²⁸⁷ No obstante, los niños con edades menores a los tres años no manifiestan indignación y deseo de represalia ante los comportamientos ajenos aprovechados, lo cual admite suponer que el cerebro de tres años ha desarrollado dispositivos neurofisiológicos aptos para la interacción social que aún no existen en edades más tempranas. Antes de los dos años de edad los niños pequeños ya han desarrollado la capacidad de empatizar y la motivación altruista, pero todavía no son capaces de comprender la intencionalidad ajena y sentir emociones de repulsa ante los comportamientos de explotación (Warneken y Tomasello 2006).

²⁸⁸ Véanse al respecto Clutier, *et. al.* (2011) y Ramachandran y Blakeslee (1999: 208 y ss.)

los desconfiados para predecir con éxito el comportamiento ajeno y evitar ser explotados (Kikuchi, *et. al.* 1996).

Estos hallazgos permiten inferir que existe lo que se puede llamar un “contrato social” epigenético en la arquitectura (moral) de la mente humana (Greene 2005; Hauser 2006). La psicología humana está adaptada para detectar incoherencias e incumplimientos de las reglas epigenéticas de condicionalidad en el intercambio social, lo cual contribuye al éxito de una estrategia evolutiva de cooperación basada en el reconocimiento de violaciones de la condicionalidad²⁸⁹. Además, existen pruebas de expresiones faciales y otros rasgos fenotípicos, como la expresión corporal, el tono de voz y su modulación, o la sudoración y el color de la tez, que son manifestaciones filogenéticas inconscientes de la intencionalidad y actitud relacional muy difíciles de simular, codificadas tanto para ser producidas como detectadas²⁹⁰. Los individuos con patologías neurológicas incapacitantes para la racionalidad condicional y la inferencia analógica se muestran torpes e indiferentes en la función de detectar actitudes tramposas y aprovechadas en su entorno. Esto sugiere que existe una neurofisiología evolucionada que soporta la mente social y las capacidades para participar en procesos de intercambio

²⁸⁹ La teoría está formulada en Cosmides y Tooby (1992). Un resumen de los resultados experimentales con razonamiento condicional apoyada en la metodología Wason puede encontrarse en Boyd y Silk (2004: 529-532). En Fiddick y Erlich (2010) puede hallarse un repaso y puesta al día de la literatura más relevante acerca de los estudios experimentales con la teoría de Cosmides y Tooby y el test de Wason. Asimismo presentan un experimento confirmatorio de que la cognición humana está más capacitada para la identificación de tramposos y explotadores que para la identificación de altruistas. Oda y Nakajima (2010) muestran experimentalmente una superior capacidad de identificación y memorización de los rostros de defectores que los de los altruistas. Oda (1997) ha detectado un sesgo de reconocimiento de los defectores en un experimento con el juego del dilema del prisionero especialmente en el caso de varones. Esto contribuye a reforzar y validar la tesis de la reputación como proceso mental de regulación social, especialmente teniendo en cuenta que las identidades de los individuos con reputación de defectores activan emociones desagradables, son observados con mayor atención y recordados mejor y durante más tiempo que los cooperadores (Singer, *et. al.* 2004b; Chiappe, *et. al.* 2004).

²⁹⁰ En el estudio de Fetchenhauer, *et. al.* (2010) se estima que la mayoría de los individuos predicen correctamente la disposición estable y no fingida al comportamiento altruista de otros individuos tras una breve observación visual (de 20 segundos) de su expresividad durante la acción de compartir dinero. La mente humana está provista de dispositivos especializados para detectar de manera rápida y automática intenciones y estados emocionales de otras personas, y atribuir niveles de confianza (Engell, *et. al.* 2007; Rudoy y Paller 2009; Todorov, *et. al.* 2007). Asimismo, se ha demostrado experimentalmente que la atribución de intencionalidad en la acción observada, ya sea en actos de ayuda o de estorbo, mejora la memorización de la identidad del actor (Camilleri, *et. al.* 2010). Existen expresiones faciales de estados emocionales no controlables conscientemente que señalizan la sinceridad o insinceridad del altruismo empírico, así como el carácter altruista o no de las personas (Mealey, *et. al.* 1996; Schug, *et. al.* 2010). Véase al respecto de la anatomía gestual involuntaria de la musculatura facial, y su captación por parte de los participantes en experimentos de visualización, los trabajos de Ekman (2003), Brown y Moore (2003), y Ten Brinke, *et. al.* (2012).

equitativo. La operativa se lleva cabo a través de la cognición de las intenciones ajenas mediante la interpretación inconsciente de la mímica facial, gestos corporales, y otra comunicación verbal y no verbal (Stone, *et. al.* 2002). El sistema completo viene reforzado por las motivaciones de reciprocidad y gratitud, que inducen a experimentar sensaciones agradables en los cooperadores cuando se percibe un ambiente de cooperación (Rilling, *et. al.* 2002). Asimismo, la experimentación muestra una clara tendencia a incrementar la tasa de cooperación por parte de los individuos con aquellos de los que se tiene conocimiento sobre un historial cooperativo y/o expectativas de que van a cooperar (Barclay 2004). En definitiva, durante la filogénesis de una especie social que depende de la sociedad para sobrevivir, como la humana, no parece desacertado aceptar que han evolucionado adaptaciones para preservar la cooperación y el orden social (Clutton-Brock y Parker 1995).

A todo esto cabe añadir que en ambientes sociales en que el sujeto presume o comprueba una atmósfera de cooperación se tiende a cooperar en mayor grado, a reforzarse el sentimiento de pertenencia y a mostrar una superior generosidad, así como a la inversa²⁹¹. Las estructuras de acción colectiva en que resulta posible una mejor vigilancia de las posibles conductas oportunistas, y la aplicación de castigos a los defectores localizados, registran tasas de cooperación más elevadas, y viceversa (Fehr y Gächter 2000a; Frey y Meier 2004b; Kurzban, *et. al.* 2001; Sausgruber 2009). Ello no quiere decir que los patrones de conducta sean automáticos y deterministas, sin una previa evaluación cognitiva. Las creencias acerca de la disposición a cooperar del cojugador influyen en el grado de cooperación propio. Así, aun disponiendo de un perfil conductual proclive al altruismo, los cooperadores manifiestan flexibilidad de comportamiento, que incluye tanto la defeción como incluso la conducta de castigo malintencionado, en función de sus expectativas sobre los demás (Charness y Rabin 2002; Saijo y Nakamura 1995). Una proporción nada desdeñable de sujetos recompensan con reciprocidad positiva a los que cooperan, e imparten castigos a los que se aprovechan mediante conductas no cooperativas; y lo hacen incluso soportando un coste personal por ello (Fehr y Gächter 2002; Fehr y Fischbacher 2003).

²⁹¹ Sobre la influencia del ambiente en la disposición conductual a cooperar existe una amplia literatura, entre la que se puede restaltar Traulsen, *et. al.* (2010); Leider, *et. al.* (2009); Fischbacher, *et. al.* (2001); De Cremer y Leonardelli (2003).

En el plano neurofisiológico, especialmente en los machos de la especie humana se inhibe significativamente la respuesta empática cuando aprecian el sufrimiento de personas no cooperativas. Este proceso ocurre en paralelo a la activación incrementada del sistema neurológico de recompensa como expresión de la satisfacción del deseo de venganza. La cooperación estimula una respuesta neurológica empática y de atracción, mientras que la defección estimula la reacción de venganza y exclusión (Singer, *et. al.* 2006; Seymour, *et. al.* 2007). Igualmente, la experimentación neurocientífica muestra que las creencias adquiridas acerca de la catadura moral de otro individuo influyen de manera efectiva en las expectativas sobre él, en la confianza que se le deposita, y en el grado de preferencia a cooperar con él (Delgado, *et. al.* 2005)²⁹². Dado que existen procesos cognitivos de alerta implicados en la psicología de la confianza, la reputación se muestra pues como una información altamente relevante en la motivación del comportamiento cooperativo. Del mismo modo, los individuos revelan fehacientemente preocupación por exponer públicamente comportamientos que deterioren su reputación, y aumentan su nivel de cooperación cuando se presumen o saben observados (Mittone y Ploner 2011).

Para finalizar, conviene concretar que la especie revela además una enfática preferencia por los cooperadores de calidad y por cooperar con ellos; asimismo por eludir las relaciones con los aprovechados y explotadores, y disfrutar con la desgracia de estos últimos. Si la cooperación y el altruismo no son conductas indiscriminadas sino que, como se ha demostrado, se trata de conductas selectivas que refuerzan la eficacia inclusiva de caracteres (*assortative interaction*), entonces resulta factible encuadrar la evolución de la cooperación en procesos de selección en varios niveles. Ello se explica porque las redes de cooperadores obtendrán mejores resultados de eficacia que los individuos solitarios, las redes de egoístas, o los entornos relacionales mixtos e indeterminados (Sethi y Somathan 2003).

²⁹² Ha de tenerse en cuenta que la psicología humana manifiesta una clara preferencia por relaciones sociales conducentes a procesos de selección inclusiva. Las capacidades de selección y discriminación están relacionadas con la preferencia empírica por la semejanza. Así, se ha comprobado que los individuos con una predisposición conductual relativamente más altruista (egoísta) tienden a relacionarse con otros igualmente de predisposición más altruista (egoísta), y a establecer conexiones, redes y vinculaciones sociales de amistad con otros altruistas (egoístas) (Leider, *et. al.* 2009).

Los estudios de Leider, *et. al.* (2009, y 2010) sobre redes sociales humanas concluyen en la necesidad de más investigación en orden a conocer si el rasgo altruismo es exógeno (replicación genética y reconocimiento de señales altruistas) o endógeno (replicación cultural dentro de redes conformadas aleatoriamente). Según Leider, *et. al.* (2010) las pruebas corroboran un déficit del modelo de señalización y reconocimiento filogenético del rasgo altruismo, aunque se demuestra una superior generosidad empírica (del 50%) en las relaciones de altruismo entre amigos respecto a las producidas con desconocidos. Esto confirmaría la hipótesis de que las preferencias altruistas se replican culturalmente a través de la interacción social.

Dado que la tesis del altruismo filogenético tiene una demostrada solvencia científica, queda por analizar el papel jugado por la cultura y la evolución cultural en la ontogénesis de la cooperación y el altruismo. La cultura puede ser un factor de refuerzo o inhibición de disposiciones innatas. En este aspecto conviene profundizar sobre los marcadores culturales, que permiten señalar e identificar disposiciones de cooperación en las relaciones sociales. A ello dedicamos el siguiente capítulo.

4.7.- Conclusiones

En el presente capítulo se ha analizado teóricamente la evidencia empírica del altruismo exogámico y exodémico, y se han aportado pruebas que permiten aceptar las teorías propuestas. La especie humana se halla especialmente adaptada dentro de la naturaleza para cooperar con desconocidos, incluso en ambientes de anonimato y en interacciones de única jugada. Esto permite proclamar una naturaleza humana en la que se encuentran presentes motivaciones para el comportamiento altruista, y un conjunto de adaptaciones determinantes de una psicología capacitada para el altruismo.

No es posible comprender el comportamiento altruista fuera de un marco evolutivo y sin reconocer cómo la evolución ha provisto de adaptaciones específicas que se manifiestan en la complejidad del individuo empírico, por encima de reduccionismos teóricos que son poderosos en simplicidad pero vulnerables en capacidad explicativa. El comportamiento altruista no puede explicarse solamente desde la perspectiva del individualismo metodológico y además, lejos de admitir únicamente

alguna/as de las teorías evolutivas convencionales, como es el caso de la selección de parentesco, la reciprocidad y la selección de grupo, aquí se propone que es posible explicar la evolución del comportamiento altruista desde una perspectiva que integra la selección individual con que ésta ocurra en ámbitos orgánicos y relacionales que, de manera efectiva, alinean los intereses evolutivos del individuo con aquéllos de otros individuos con quienes se comparte un *pool* de rasgos replicables.

El altruista, de este modo, puede ganar eficacia con su comportamiento porque existen adaptaciones orgánicas que refuerzan el comportamiento altruista mejorando la homeostasis, y por lo tanto la salud y calidad de vida del individuo altruista. Adicionalmente, otras adaptaciones filogenéticas como los sentimientos morales de la admiración y la gratitud, refuerzan socialmente el altruismo a través de dispositivos neuroendocrinos motivadores de la reciprocidad positiva (directa e indirecta) hacia aquéllos individuos que señalizan convincentemente su calidad altruista. Algunas de estas adaptaciones se expresan epigenéticamente en regulaciones institucionales que conforman órdenes sociales en los que se premia la excelencia prosocial. Es el caso de los sistemas culturales de prestigio, que recompensan la eficacia perdida por el altruista al realizar comportamientos prosociales.

¿Cómo se puede explicar la evolución de estas adaptaciones motivacionales que capacitan para el altruismo, así como de los procesos orgánicos y psicosociales de refuerzo de las mismas? Da la impresión de que los órdenes natural y social que hacen posible la evolución del comportamiento altruista proceden de adaptaciones coevolutivas determinantes de una suerte de «coreografía oculta» (Bowles y Gintis 2008), o «mano invisible moral» (Richerson y Boyd 2008). En este capítulo se han analizado dos fuerzas evolutivas, o sistemas ecológicos de causación última, capaces de explicar el contenido de esa metafórica coreografía oculta. Estas fuerzas sustentan dos teorías explicativas del altruismo interrelacionadas y complementarias, que son: (a) la teoría del destino común y, (b) la teoría del equilibrio poblacional. Ambas se fundamentan en la premisa apuntada por Gazzaniga (2006: 174) consistente en que: «La evolución salva al grupo, no solo a la persona, porque parece que la salvación del grupo salva también a la persona».

En el s. XVII Spinoza (*Et: IV, xviii, e*) proclamó un acertado elogio de la cooperación consistente en que cooperar con aquellos que cooperan es la mejor decisión que se puede tomar siguiendo la guía de una razón natural que hoy día podemos denominar racionalidad ecológica de fondo. Esta instancia a cooperar con cooperadores implica que cooperar siempre es más provechoso que no hacerlo, luego como hay que cooperar por sus inherentes ventajas de supervivencia, se hace preciso identificar a los más afines y seleccionar aquéllos caracterizados por su excelencia cooperativa. Si la capacidad de cooperar es un rasgo señalizable e identificable por los individuos que son afines en la cooperación, el verdadero altruismo es más eficiente que el falso altruismo o la defección, puesto que la cooperación entre cooperadores es más eficiente para el cooperador que cualquier otra alternativa. Además, los cooperadores de alta calidad gozan de la reputación de ser buenos socios para cooperar, lo cual les permite extender el rango de posibles interrelaciones de cooperación y seleccionar los socios igualmente de mejor reputación. Esto implica que la virtud (la calidad prosocial) recibe su correspondiente premio, tal y como expresó Spinoza, aunque teóricamente, según las visiones utilitaristas no evolutivas, pueda considerarse irracional ser virtuoso.

La coreografía oculta de esta disposición a la calidad prosocial se explica, dentro de una perspectiva evolucionista, en que la participación en procesos de acción colectiva para el logro de objetivos comunes selecciona sólo a los individuos que portan rasgos de cooperación siempre y cuando la participación les sea necesaria, y el beneficio derivado de la ganancia asociada al logro común supere el coste esperado del mismo, así como la ganancia esperada de la defección. Cuando esto es así, en efecto, como apunta Gazzaniga la evolución salva al grupo porque salvando al grupo se salva el individuo. Este es el factor de evolución que subyace tanto a la teoría del destino común como a la perspectiva poblacional. De este modo, cuando la supervivencia del individuo depende del logro de metas comunes, resulta ecológicamente racional que los individuos busquen replicadores semejantes, y no porque pretendan deliberadamente la eficacia inclusiva como causa última directa de su búsqueda, sino porque con los replicadores afines es posible una cooperación más certera, rica y eficiente; esto es, una mayor eficacia de los replicadores que constituyen la propia identidad orgánica y psicológica.

La teoría del destino común se basa en que cuando dos o más agentes persiguen un objetivo común que requiere su participación respectiva, se produce un alineamiento espontáneo de intereses incluso aunque los agentes tengan una racionalidad individualista y egoísta (p.ej. propulsar un tándem cuyo movimiento necesita una fuerza superior a la de un solo ciclista). En estos casos los individuos actúan como si de un grupo se tratara, aunque no tengan conciencia de ser un grupo delimitado por identidad colectiva (biológica o cultural) alguna, ni conozcan personalmente a los miembros de su “equipo”. Ello se explica porque la cooperación aporta ventajas adaptativas a los individuos que desarrollan la capacidad de instrumentar estrategias de cooperación (procedimentales, y en casos de mayor complejidad, incluso psicológicas). La cooperación será pues una adaptación prosocial que favorecerá a los individuos que disponen de ese rasgo como cualquier otro rasgo con valor de supervivencia.

Lo que compromete de base a los individuos con la cooperación en pos de los objetivos conjuntos no es *a priori* un sentimiento de vinculación ni tampoco la reciprocidad, sino su destino común, su dependencia mutua. En estos casos los agentes cooperan aunque no se hayan planteado intencionalmente el hacerlo. No les queda más remedio. Si dos agentes egoístas que persiguen un fin común e indivisible se conocen mutuamente, entonces la cooperación será intencional y funcionarán conductualmente como un equipo cognitivamente delimitado. La toma de decisiones se realizará con una racionalidad de equipo, y la información que se transmitan será utilizada para promover los intereses comunes. Como explica Frank (2008b: 274-275):

Si en una transacción coinciden completamente los intereses de las dos partes, éstas no tienen motivos para engañarse una a la otra. Cuando una jugadora de bridge utiliza las señales convencionales para informar a su compañera de que tiene una buena baza, basta con que dicha información sea inteligible. No hay razón para que su compañera dude de su sinceridad.

Lo relevante es el grado de comunidad en los objetivos individuales, eso es lo que establece el replicador y lo que determina el comportamiento cooperativo. Y esto nos conduce a la evolución de causas próximas que posibilitan decidir desde los intereses de una perspectiva poblacional. En efecto, si los individuos fueran vehículos portadores de algo común, y su función objetivo consistiera en maximizar la supervivencia y

propagación de ese “algo”, la elección racional sería cooperar; y si se trata de organismos conscientes capaces de reconocer en otros ese “algo”, entonces su utilidad incluiría la utilidad de aquellos con los que se identifican, es decir, actuarían con una racionalidad de grupo en la que la cooperación es el rasgo conductual dominante²⁹³. En organismos psicológicos, como el ser humano, no sería nada extraño que la evolución hubiera implantado además instintos y facultades cognitivas que capacitaran a los sujetos para preocuparse por el bienestar y la salud de los otros, y por sus intereses particulares coherentes con el objetivo común. Al contrario ocurriría con los instintos y facultades discriminadoras adaptados para eludir y combatir la existencia y la acción de los individuos cuyos intereses se oponen al objetivo común en el que se empeña el individuo. Esos intereses individuales coherentes con alguna meta común son de hecho los intereses comunes del grupo (el “interés general”), y aquellos intereses individuales que contribuyen de manera más eficiente al objetivo común se impondrán, por racionalidad ecológica, sobre los menos eficientes.

Así pues, en la evolución de la cooperación lo relevante no es el replicador “cooperación”, o sea, no se trata de la codificación del comportamiento cooperativo en replicadores ya sean genéticos o culturales. Lo que ha permitido la evolución de la cooperación es la codificación de la capacidad de asumir objetivos comunes, que pueden ser individualistas (cooperación por conveniencia, reciprocidad), o grupalistas (p.ej, los genes o implantes culturales comunes que porta cada individuo) (cooperación por afinidad, o altruismo fuerte). De ahí la posibilidad de evolución de dispositivos adaptados a la selección multinivel. Es decir, el mecanismo de selección que se activa en cada circunstancia depende de si el objetivo compartido es o no un objetivo verdaderamente transversal. En casos de objetivos transversales la evolución puede haber implantado un sistema no egocéntrico de cognición y de toma de decisiones, una metafórica función de utilidad transpersonal, que se traduce experiencialmente en una visión del mundo no desde los ojos de un sujeto egocéntrico que persigue únicamente sus objetivos y desconoce los de otros individuos, sino desde la perspectiva de un agregado que trasciende al individuo y que de alguna manera actúa a través de él: una

²⁹³ Según esta hipótesis diversos innatismos que motivan el comportamiento, como los instintos afiliativos que en el texto hemos llamado instinto de preocupación por el bien de la comunidad, instinto de entusiasmo militante, necesidad de pertenencia, identificación progrupal, etc., forman parte del dispositivo evolucionado para la racionalidad plural.

familia, un grupo cultural, un mito, una divinidad, un héroe, una idea, una teoría, etc., o sea, una perspectiva integradora que soporta emotiva y cognitivamente una determinada identidad grupal.

La distintiva disposición a los comportamientos progrupales exogámicos y exodémicos de la especie humana es posible porque la naturaleza humana se desenvuelve en un nicho cognitivo, y por lo tanto se trata de una naturaleza cultural. Esto supone la capacidad de establecer instituciones culturales de refuerzo de la cooperación que faciliten la identificación de destinos comunes -y por lo tanto replicadores comunes- de la que aflora una racionalidad grupal. Es el caso de regulaciones y marcadores de identidad creados por el ser humano y transmisibles por herencia cultural, así como estructuras y ordenamientos sociales que sustentan patrones de acción colectiva favorables a la evolución de la cooperación.

En los siguientes capítulos se aborda el estudio de los procesos de causación próxima que explican el comportamiento cooperativo humano sobre la base de una identidad grupal (capítulo 5), así como aquellos que explican el fenómeno de la intersubjetividad (capítulo 6), una capacidad psicológica necesaria y suficiente para explicar el comportamiento humano dirigido por una racionalidad grupal.

Introducción a los capítulos 5 y 6

Conexión de vida y psiquismo: replicadores culturales e identidad humana

El capítulo precedente finalizó con una reflexión acerca de los estudios de Leider, *et. al.* (2009, y 2010) sobre redes sociales humanas. Estas investigaciones concluyen que no es posible establecer con certeza si el rasgo altruismo es un rasgo filogenético exógeno, como es la morfología humana o, por el contrario, se trata de un rasgo ontogenético endógeno, adquirido por replicación cultural bien sea de forma independiente, o bien asociado a cualquier otro rasgo filo u ontogenético. Según los hallazgos experimentales de Leider y sus colaboradores la supuesta propensión empírica a un comportamiento altruista aleatorio, a través de la señalización y el reconocimiento de señales reveladoras de la veracidad genética del rasgo, no es suficientemente robusta en comparación con la detectada en redes sociales de amistad y afinidad cultural en general. Además del reducido ámbito del altruismo familiar, el altruismo humano se expande hacia relaciones de afinidad cultural, por lo que se trataría de un rasgo de manifestación selectiva y circunscrito a determinados marcadores ontogenéticos de identidad. Este puede ser el caso del “contagio de afinidad” que surge del conocimiento relacional previo, o la identificación a través de algún símbolo adoptado culturalmente que informe acerca de una inequívoca identidad de pertenencia al grupo. La identidad grupal presupone ante todo una disposición a cooperar con aquellos comprometidos a cooperar selectivamente y sobre la base de algún identificador de afiliación.

Por otro lado, dentro de la sociobiología experimental está siendo discutida la tesis del haplodiploidismo endogámico de las agrupaciones de himenópteros, entendido como única explicación del comportamiento altruista en estas especies. Esto se debe a que hay constatación empírica de que el demarcador de la identidad de los individuos que pertenecen a una colmena u hormiguero no es de origen genotípico (algún rasgo codificado genéticamente que comparten únicamente los parientes), sino fenotípico (estar impregnados de la feromona regia). Una prueba de ello es que el coeficiente de relación genética empírico que hay en las colmenas de abejas melíferas es en ocasiones

significativamente inferior al 0,75 teórico. La causa está en que con frecuencia las reinas son fecundadas por más de un macho, y de ello nacen crías de padres distintos. Consecuentemente, conviven individuos en la misma población con diferentes niveles de parentesco; sin embargo, no por ello se aprecian comportamientos discriminatorios o agresivos entre los individuos menos emparentados (Dugatkin 2007: 185-186). Como en la especie humana, el elemento demarcador del ámbito social de la cooperación tiene una explicación más amplia que el estricto parentesco, en este caso compartir una misma feromona que impregna la colonia. Un individuo con un coeficiente de parentesco de 0,75 que no esté impregnado de la feromona regia es atacado hasta la muerte por sus hermanas genéticas. Así pues, la cultura y la evolución cultural, ejemplificadas como influencia del ambiente social en el individuo, pueden explicar la ontogénesis de la cooperación y el altruismo tanto exogámico como exodémico²⁹⁴.

La demostración empírica de la existencia de sesgos prosociales a favor de ciertos rasgos fenotípicos de origen cultural, como por ejemplo la cooperación preferente hacia los individuos que pronuncian un acento similar (Cohen 2012), o que en su identidad civil comparten nombre y/o apellidos (Oates y Wilson 2002), ha dado pie al estudio del papel desempeñado por el fenotipo, y más concretamente el fenotipo adquirido culturalmente, en la evolución de la cooperación. Existe evidencia empírica de la preferencia por la semejanza de rasgos culturales, y de que ésta contribuye a cohesionar los grupos humanos favoreciendo la evolución del rasgo “favoritismo por lo semejante” (*in-group favoritism*) (Efferson, *et. al.* 2008). Cierta grado de identificación con la simbología, ideología y propósito de los grupos humanos es preciso para que los individuos expresen una cooperación efectiva. Boulding (1975[1953]: 39) equipara este aspecto cultural a una metafórica «gravitación social».

Desde la perspectiva neodarwinista ortodoxa se trata de encontrar correlaciones genéticas en sentido amplio y no tanto genealogías contrastadas (como p.ej. la relativa similaridad en los marcadores de microsátélites del ADN), de las que se puedan inferir parentescos ocultos y adaptaciones de reconocimiento de similaridad más amplias y complejas que las conocidas hasta ahora (Dickinson y Koenig 2003: 1.888).

²⁹⁴ Véanse los trabajos de Nowak, *et. al.* (2010) y Wilson y Hölldobler (2005) sobre una revisión teórica del fenómeno de la eusocialidad.

Dado que la estricta teoría de las etiquetas fenotípicas de Dawkins, reveladoras de genotipos unívocos deterministamente altruistas, plantea no menos dificultades empíricas que la teoría de la selección de grupo, la nueva generación de modelos de eficacia inclusiva escinde el altruismo del parentesco genético²⁹⁵. En modelos como el de Colman, *et. al.* (2012) la cooperación puede evolucionar sin necesidad de una elevada correlación genética del tipo *kin selection*: es suficiente con que un mismo gen o grupo de genes produzca en los organismos la expresión fenotípica de rasgos variables, así como la predisposición a cooperar preferentemente con los organismos en los que se reconoce dicha expresión fenotípica. Así pues, el altruismo puede explicarse como una consecuencia *ex post facto* de otros replicadores implicados en la evolución orgánica, y no necesariamente como un rasgo replicable en sí mismo²⁹⁶.

Las posturas neolamarckianas contemporáneas suponen un desafío al aparente determinismo genetista de la psicología evolutiva y el neodarwinismo ortodoxos. Se está conociendo cómo el ambiente en que se desarrolla la ontogénesis resulta decisivo en la explicación de los fenotipos empíricos, especialmente en aquellas especies dotadas de una evolucionada plasticidad fenotípica²⁹⁷. Wheeler y Cark (2008) advierten sobre lo incompleto de las perspectivas centradas únicamente en los replicadores genéticos, y proponen un modelo de comprensión de la naturaleza humana compuesto por la integración, en el vehículo orgánico, de replicadores culturales que generan nichos cognitivos a través de la adaptación epigenética de módulos neurológicos adaptados por la neuroplasticidad en la ontogénesis. Esto a su vez genera procesos de creación social de condiciones de supervivencia (materiales y psicológicas) que a largo plazo influyen en el ADN de esas poblaciones por efectos trinquete y coevolución. Este modelo de “triple hélice”, que se basa en las teorías de R. C. Lewontin explicitadas en su libro *The triple helix: Gene, organism and environment* (2000), ensambla la naturaleza humana en complejas relaciones de interdeterminación (*bootstrapping*) producidas entre

²⁹⁵ Si bien la evolución del altruismo entre individuos que discriminan a favor de otros con rasgos fenotípicos similares (*similarity discrimination*) puede dar lugar a procesos de selección inclusiva de parientes en el largo plazo.

²⁹⁶ Por esta razón Gintis (2003) plantea la metáfora del altruismo como autoestopista (*hitchhiker's effect*) que “viaja” en la expresión de otros genotipos autónomos.

²⁹⁷ El nuevo lamarckismo científico enfatiza la relevancia del ambiente en la configuración ontogenética de los organismos, y lo hace sin necesidad de recurrir a la concepción tradicional del lamarckismo en lo concerniente a la heredabilidad de los rasgos adquiridos. Para el nuevo lamarckismo la epigénesis se concibe como adaptación fenotípica. Véanse Moreno (2008: caps. 9 y 13), y Oyama (2000).

replicadores simbólicos exocerebrales, los replicadores biológicos de la genética, y las adaptaciones epigenéticas de los fenotipos en replicadores orgánicos causados por la plasticidad fenotípica. Estos últimos están constituidos por estructuras orgánicas flexibles y adaptativas, como el caso de las estructura neuronales, susceptibles de replicación y surgidas de la interacción entre los replicadores genéticos y ambientales:

Estas estructuras físicas se combinan con prácticas apropiadas culturalmente transmitidas para transformar la resolución de problemas, y (en los casos más radicales) para hacer posible nuevas formas completas de pensamiento y razonamiento. (Wheeler y Clark 2008: 3.564).

El enfoque teórico del revisionismo neolamarckiano se basa en modelos de herencia epigenética, que plantean una teoría de la herencia a través de sistemas de copia de replicadores separados de los sistemas de transmisión de información puramente genéticos, como la codificación que establece en el interior del organismo la recepción de información procedente del medio externo (*host imprinting systems*):

Un niño cuyas primeras experiencias son modeladas por ambientes especiales caracterizados por los libros y los programas de ordenador, y cuyo propio perfil cognitivo emergente favorezca ciertos elementos dentro de ese nexo culturalmente posibilitado, desarrollará un sistema cognitivo que no es sólo superficialmente, sino profundamente diferente del que dispone otro niño enculturado en un ambiente distinto. (Wheeler y Clark 2008: 3.571-3.572).

En la especie humana el principal factor determinante de la epigénesis de rasgos que requieren la interacción entre genética y ambiente es la cultura. De manera simplificada puede admitirse que herencia cultural es todo aquello que constituye al individuo y que no forma parte de su herencia genética; o sea, lo que adquiere durante su vida mediante diversos procedimientos de transmisión y recepción de información y otros recursos. Obviamente, esta herencia cultural se transmite *à la* Lamarck, lo cual implica que toda esa información no se replicará a través del ADN depositado en el núcleo de las células sexuales, sino que se replicará a través de transmisión y absorción cultural que penetra por los sentidos externos. Por cultura -y por tanto por información cultural, o memes- se puede entender un espectro amplísimo de recursos informativos y bienes materiales.

Hayek (2007[1962]) consideró el ámbito de la cultura aquel que se ocupa de los modelos de comportamiento y modelos epistemológicos que son aprehendidos a través de la pertenencia a grupos y la interacción social. Lumsden y Wilson (1991: 3) definen la cultura genéricamente como «la suma total de constructos mentales y comportamientos, que incluyendo la construcción y utilización de artefactos, se transmite de una generación a la siguiente por aprendizaje social». Richerson y Boyd (2005: 5) consideran cultura toda información capaz de influir en el comportamiento de los individuos, transmitida entre los miembros de su especie a través de procesos (conscientes e inconscientes) de aprendizaje social. Entre la información transmitida que constituye la cultura se citan las ideas, conocimientos, creencias, valores, técnicas y actitudes. Ayala (1994: 225) incluye los hábitos adquiridos y estilos de vida, artes y técnicas para hacer y usar objetos materiales, el lenguaje, instituciones sociales y políticas, tradiciones éticas y religiosas, conocimientos científicos y humanísticos, etc.

John G. Bruhn (2005) asimila la cultura a un paquete implantado, una programación grabada en las mentes de los individuos, que ordena los circuitos innatos y los adapta ontogénicamente. Lo significativo a efectos fenotípicos es que tales programaciones afectan al comportamiento, y además en muchas ocasiones también comportan marcadores simbólicos perceptibles o inferibles que señalizan esos patrones de comportamiento implantados. Desde este punto de vista la asimilación cultural ordena las estructuras orgánicas y bioquímicas estableciendo constricciones cognitivas y emocionales, que a su vez influyen en el comportamiento de una forma predecible. Esta ordenación epigenética produce en la psicología humana una categorización de la conciencia de sí mismo, así como del mundo, de acuerdo a estereotipos asimilados del medio ambiente social:

Tales estereotipos no son más que categorías cognitivas que suelen satisfacer ciertas necesidades emocionales, que se muestran bastante resistentes al cambio por disconformidad informativa, y que operan como poderosos imanes cognitivos a través de los cuales la información es asimilada. (Jones y Colman 1996: 844).

El estereotipo heredado culturalmente consiste pues en una generalización estable y sobreesimplificada del sujeto y de la realidad circundante, que combinadas definen la identidad del individuo y su autoadscripción relacional respecto a otros individuos. Se

trata de creencias que llevan implícitas conexiones emotivo-cognitivas, de modo que están asociadas a juicios de valor de aceptación-no aceptación, atracción o rechazo.

La herencia cultural interviene en la construcción semántica de una visión del mundo, del sentido de la existencia y del sí mismo. Se trata de la herencia de una teoría metafísica de la realidad, un “paradigma” en el lenguaje de la epistemología kuhniana: «Los paradigmas estructuran la observación y, por lo tanto, definen la realidad» (Ruse 2001: 37). La realidad, y las teorías explicativas de la misma; las técnicas, herramientas, creencias y visiones, etc., esto es, la cultura ampliamente entendida, son fruto de complejos procesos de construcción social, de interacción de mentes que contribuye a focalizar la atención, denominar y reconocer los conceptos y las emociones, y a seleccionar las herramientas de observación que son aceptadas. Y esto es así, porque frente al realismo popperiano, «no hay conocimiento sin conocedor, y el modo en que llegamos a aceptar un nuevo paradigma puede depender de manera crucial de nuestra propia historia» (Ruse 2001: 38). Citando a Stephen J. Gould,

Los hechos no son retazos puros e inmaculados de información; la cultura también influye en lo que vemos y en cómo lo vemos (...) Las teorías no son inducciones inexorables a partir de hechos. Las teorías más creativas son con frecuencia visiones imaginativas que se imponen sobre los hechos; la fuente de la imaginación también es en gran medida cultural. (Citado en Ruse 2001: 45-47).

Se puede entonces hablar de marcadores culturales que delimitan la comprensión y el sentimiento del individuo ante sí mismo y ante el mundo, y que además lo identifican respecto a los demás individuos. Estos marcadores se propagan en la población a través de la interacción social y no siempre de forma consciente. Pero, ¿dónde está la cultura?, ¿cuál es su soporte físico en el cerebro?, ¿cuál es la unidad de análisis del replicador cultural? Prescindiendo de los soportes físicos y simbólicos de transmisión cultural, como la imagen, el sonido o la escritura, ¿cómo se almacena la cultura en los cerebros?

Una importante tradición de la antropología, de enfoque no evolucionista, establece que la cultura es un fenómeno emergente que surge de la interacción social, y que no se halla en el interior de los organismos (Sperber 2001). Según este enfoque, los *engramas* que conforman la “mente cultural” no están dentro de los cerebros sino fuera

de ellos y, lo que resulta más insólito, la propia “mente” no es un producto del cerebro individual, sino un fenómeno complejo de tipo exocerebral y relacional, una especie de meta-mente colectiva constructora de una realidad socialmente construida y compartida²⁹⁸. Esta explicación holística confiere sentido a la conocida tesis del antropólogo Clifford Geertz sobre que las culturas no están *en* las mentes de los individuos, sino *entre* ellas, dentro de un sistema social de conocimiento al que están adaptadas las mentes individuales.

¿Dónde está pues la unidad de análisis de la cultura?, ¿es posible asegurar *a priori* que un cerebro-mente individual es portador de algún tipo de rasgo representativo de una determinada cultura? Castro Nogueira y Toro Ibáñez (2006) afirman al respecto que «la cultura es un entramado de caracteres continuos que se mezclan y cambian, donde difícilmente pueden distinguirse las características de un replicador».

En las últimas décadas se han producido avances teóricos a la hora de plantear formalmente la transmisión cultural, con resultados analíticos muy relevantes²⁹⁹. Sin embargo, la contrastación de los modelos ha tenido un éxito relativamente menor, salvo conocidas excepciones como p.ej. Cavalli-Sforza, *et. al.* (1982). Esto se debe en buena medida a la dificultad de concebir la unidad empírica de análisis, o la “partícula cultural”, que se transmite y en la se producen variaciones dentro de procesos de evolución. Consecuentemente, los estudios de la evolución cultural, que se basan en apriorismos teóricos más que en la definición y clasificación de unidades empíricas de selección, son acusados de -a diferencia de la genética de poblaciones- carecer de rigor científico o no ser contrastables. No obstante, como plantean Boyd, *et. al.* (1997), la filogénesis cultural es teóricamente concebible y puede ofrecer resultados explicativos, si bien el replicador cultural, hoy por hoy, puede parecer un concepto cuasi-metafísico poco concluyente y difícil de manejar empíricamente. Se necesitarán avances ulteriores en las neurociencias, si es que fueran posibles, para demostrar cómo se establecen los memes en los organismos, cómo condicionan su conducta, y cómo se manifiestan fenotípicamente; esto es, cómo se desenvuelven los replicadores fenotípicos en la dinámica evolutiva del mundo real a través de su integración orgánica.

²⁹⁸ Véase al respecto el trabajo paradigmático de Berger y Luckmann (1986[1966]). Un trabajo más reciente es el de Fischer (1990).

²⁹⁹ Una panorámica de la literatura más relevante puede encontrarse en Bisin y Verdier (2008).

En todo este sentido, el aspecto relevante de la cultura en la explicación evolutiva de la cooperación es que los marcadores de identidad operan insertando en el individuo condicionamientos (como los estereotipos, las creencias o las costumbres) que encauzan los innatismos biológicos hacia conductas prosociales. Otra capacidad de los marcadores de identidad es la que supone funcionar como señales informativas perceptibles (o inferibles) que posibilitan la identificación de otros individuos que han asimilado los condicionamientos prosociales. Se trata pues de aprendizajes, señuelos y señales, que combinados producen un comportamiento cooperativo dentro del *pool* de replicadores de un paradigma cultural. La transmisión cultural y la generación consecuente de similitud fenotípica pueden ser variables explicativas de la evolución de la cooperación.

Si la cooperación es capaz de evolucionar a través de la preferencia por la similitud fenotípica, y los fenotipos expresados pueden ser variables, un factor de evolución (un sistema de causación última) candidato a explicar el altruismo humano es, destacadamente, las adaptaciones filogenéticas capacitantes para interiorizar cultura y expresar esa “impregnación” a través de fenotipos culturales. En el lenguaje antropológico y sociológico se está haciendo alusión a dos fenómenos:

- a) El conformismo, o receptividad a la influencia social (adoctrinamiento), tema que se estudia en el capítulo 5; y
- b) la preferencia por lo semejante (homofilia), asunto estudiado en el capítulo 6.

Ambos aspectos constituyen dos ingredientes fundamentales de la identidad psicológica humana y de los procesos de identificación interpersonal. En esta perspectiva el altruismo puede evolucionar porque incrementa diferencialmente la eficacia de los organismos que interaccionan con sus semejantes en los aspectos señalados. Se trata pues de un forma de eficacia inclusiva dentro de un grupo/población, pero sin el requisito del parentesco genético.

Los experimentos realizados muestran que la homogeneidad produce mejores resultados de cooperación que la heterogeneidad. El que se pueda asegurar un elevado grado de homogeneidad, por ejemplo, en las actitudes prosociales de los individuos que componen un grupo, influye en el incremento de la tasa de cooperación desde la primera hasta la última jugada (Ones y Putterman 2007). Está constado que la interacción entre cooperadores de alto compromiso es evolutivamente estable sin necesidad de castigos y controles externos, cuando estos tienen la certeza de que todos los miembros del grupo son cooperadores de alta calidad (Gächter y Thöni 2005). Dado que los grupos integrados de cooperadores de alta calidad generan más eficacia que el más productivo de aquéllos de inferior calidad, o los que mezclan cooperadores y defectores, se cumple un supuesto evolucionista que permite anticipar la evolución de adaptaciones que faciliten a los cooperadores el agruparse con sus afines y establecer filtros restrictivos de la intromisión de defectores.

Estos resultados demuestran que la homogeneidad del rasgo cooperación explica el éxito evolutivo del mismo. Sin embargo, lo que aquí se va a analizar es si la homogeneidad de cualquier otro rasgo como pueden ser, por ejemplo, la fuerza física o la estatura, la creencia en un dios o en una teoría científica, también contribuye a reforzar, y cómo lo hace, el éxito evolutivo de la cooperación a través de procesos indirectos. Esto quiere decir que se precisa un análisis sobre si la homogeneidad en sentido amplio (y no sólo de una tendencia genética prosocial, como el modelo de la “barba verde” de Dawkins) es una variable explicativa de la cooperación a través de la eficacia inclusiva de otros replicadores no genéticos³⁰⁰.

Los cerebros más enculturados y enculturables (adiestrados, domesticados) del reino animal son los humanos. Esta enculturación posibilita la replicación en condiciones favorables tanto de rasgos biológicos como culturales. Cabe entenderse pues que el etograma humano incluya instintos capacitantes para una enculturación atenuada de fricciones, el comportamiento moral y la sociedad, por un lado, y dado que esta normalización cultural incrementa la eficacia de los individuos, por otro lado cabe concebir otros instintos asociados a la búsqueda de los afines, la motivación para la

³⁰⁰ El modelo de señalización del rasgo genético altruismo a través de etiquetas fenotípicas se expone en Dawkins (1993[1976]: cap. 6).

simpatía e identificación con ellos y, como productos evolutivos de refuerzo, la lealtad, el vínculo y el altruismo:

Las reacciones instintivas de carácter moral se basan en capacidades evolucionadas. Una capacidad pertinente es la identificación profunda con el grupo de iguales, que constituye la base de gran parte de lo que nos hace únicos a los seres humanos (...) el instinto comunitario nos hace diferentes de los demás animales. Nos permite identificarnos con un grupo de personas más grande, marcado simbólicamente, como una tribu, una religión o una nacionalidad. La mayoría de las personas están deseando de pertenecer a un grupo social más allá de la familia, y vincularse emocionalmente al mismo. Muchos están dispuestos a vivir y morir por su grupo étnico o su religión (...) si alguien se entusiasma cuando ve jugar a su equipo, pero se siente poco estimulado para ver partidos que juegan otros equipos aunque sean mejores, entonces es que sigue su instinto tribal. Si su máximo placer es la calidad del juego más que el éxito de su equipo, es que su amor al deporte se ha librado de la identificación con la tribu. Este es el caso de pocos. (Gigerenzer 2008: 212, y 230-231).

En los próximos capítulos se estudian estas cuestiones con mayor profundidad, que se refieren a la naturaleza de las relaciones del sujeto con sus semejantes. Se trata en definitiva de explorar:

- a) Por un lado, el fenómeno de la identidad humana, dentro del capítulo 5;
- b) en segundo lugar la intersubjetividad, dentro del capítulo 6,

y la influencia de ambas en el comportamiento altruista, exoparental y exogámico, característico de la especie humana.

Capítulo 5.- La evolución cultural del altruismo: Grupalismo e identidad

... las propiedades individuales importantes para la existencia y la preservación del grupo, y por consiguiente también para la existencia y preservación de los propios individuos, han sido modeladas a través de la selección de las procedentes de los individuos que viven en grupos que en todo estadio de la evolución han mostrado tendencia a obrar basándose en reglas tales que han hecho que tales grupos fuesen más eficientes. (Hayek 2007[1967]: 123).

... nuestras mentes evolucionaron para valorar inherentemente la existencia de grupos por sí mismos, de manera que contamos con una estructura de coordinación (de las acciones) disponible y a la espera de cuando surja la necesidad. Del mismo modo que los humanos valoramos cualquier otro bien intrínseco, estamos diseñados para apreciar la existencia y la continuidad de los grupos a los que pertenecemos (...) Nuestra psicología evolucionada incluye una variable motivacional que rastrea la capacidad social del grupo o grupos de los que el individuo es miembro. Esta variable regula el interés por encontrar o desarrollar grupos si el individuo carece de una cantidad suficiente de estructuras de apoyo. *Buscamos proveernos de, crear, y perpetuar identidades grupales.* (Tooby, et. al. 2006: 111-112; paréntesis y cursivas añadidos).

El sistema de creencias y valores compartidos crea una identidad entre los miembros de la red social, identidad que rebasa el sentido de pertenencia (...) el comportamiento de la gente está moldeado y restringido por su identidad cultural que, a su vez, refuerza el sentido de pertenencia (...) Asimismo, la identidad cultural refuerza las defensas de la red al crear un perímetro de significado y expectativas que limita el acceso a ella de personas e información (...) Esta situación no es muy distinta de la de la red metabólica celular, que produce y recrea de manera continua un perímetro -la membrana celular- que la delimita y le confiere su identidad. No obstante, hay algunas diferencias cruciales entre los contornos celulares y los sociales. Los límites sociales no son necesariamente físicos, sino de significado y expectativas. Tampoco rodea, en sentido literal, la red, sino que existe en un ámbito mental que carece de las propiedades topológicas del espacio físico. (Capra 2003: 123).

La condición de sujeto comporta, al mismo tiempo que el principio de exclusión, un principio de inclusión; éste nos permite incluirnos en una comunidad, un Nosotros (pareja, familia, partido, Iglesia) e incluir este Nosotros en nuestro centro del mundo (...) El nombre de familia, de tribu, etc., instala el Nosotros en la identidad del sujeto. (Morin 2004: 79, y 79n₁).

5.1.- Introducción

En los capítulos precedentes se ha estudiado la hipótesis de evolución del comportamiento altruista dentro del enfoque analítico de la selección individual. Se concluyó verificar esta posibilidad, si bien realizando un apunte final sobre la hipótesis planteada por Leider, *et. al.* (2009, y 2010) acerca de la evolución del altruismo por replicación cultural. Esto nos sitúa dentro de un enfoque complementario de estudio, en el cual se abre paso la hipótesis de la evolución del altruismo dentro de un modelo de selección multinivel, que incluye la selección de grupo.

En el presente capítulo analizamos el papel que desempeñan los procesos de selección cultural-grupal en la evolución de la cooperación humana a través de la transmisión de identidades sociales y culturales. Específicamente atenderemos a la selección cultural como factor de evolución del comportamiento altruista. Para ello se estudia la funcionalidad de la instintiva capacidad humana de recibir adoctrinamiento, y el consecuente conformismo prosocial hacia los afines culturales (propuesta [4] de Hirschleifer según la enumeración especificada en la Introducción del capítulo 3).

Por identidad social y cultural asumimos la definición de Davis (2006: 372-373): «la identificación de un individuo *con* otros, o la idea de que el individuo se adhiere a una identidad grupal adscribiéndose a las características que pertenecen a ese grupo» (Davis 2006: 372-373. Cursivas en el original). Los procesos de generación y transmisión de identidades sociales son procesos de difusión de replicadores culturales explicativos de fenotipos. Gil-White (2001) explica que el cerebro humano está adaptado para concebir las relaciones con otros grupos culturales, tales como la etnia o la casta en las sociedades tradicionales, o el partido político, la religión o el equipo deportivo en las sociedades modernas, como si se tratara de relaciones interespecíficas. Esta confusión de fenotipos es un obvio error cognitivo, derivado de una hipotética concepción esencialista de la naturaleza. Sin embargo, la operativa puede estar basada en una exadaptación coevolutiva procultural, con valor de supervivencia, que establece una selección de identidades apoyada en procesos de compatibilidad, complementariedad y sinergias relacionales.

Existe evidencia experimental de que las identificaciones y afiliaciones culturales afectan directamente al funcionamiento de regiones especializadas del cerebro, concretamente aquellas que intervienen en la generación de experiencias de atracción – repulsión, identificación – rechazo. Asimismo, los cerebros afinan su operativa de reconocimiento y cognición en mayor o menor medida dependiendo de si a quien se percibe es o no un referente de la identidad asumida como propia (véase Kaplan, *et. al.* 2007). El cerebro expresa activaciones de procesos neurales diferentes en función de si se están realizando juicios de los estados mentales (de sí mismo y) de personas a quienes se considera semejantes o, por el contrario, de personas percibidas como diferentes en algún aspecto sustancial. En el primer caso se produce una simulación de los estados mentales ajenos sobre la base del conocimiento que se tiene de sí mismo. En el segundo caso se infieren los estados mentales de los otros en función de un modelo más abstracto de conocimiento del mundo. El fenómeno expresa la activación de patrones de cognición y relación que podemos llamar “caliente”, o identificada, con los afines, y “fría”, o desidentificada, con los que no se consideran afines (Mitchell, *et. al.* 2005). La identificación interpersonal, más probable e intensa con los considerados afines, propicia, por otra parte, la activación de motivaciones tales como la preocupación empática, la apertura a las necesidades del otro, compasión, solidaridad, el favoritismo, o simplemente comportamientos prosociales, aun en casos en que resulta posible desviar la atención o escaquearse (Batson, *et. al.* 2005; Batson 1992). Estas expresiones de la sociabilidad humana pueden ser consideradas exadaptaciones que permiten operar prosocialmente más allá de la esfera inmediata de la familia, e involucrarse en el apoyo comprometido a causas sociales abstractas, ideologías, y otros proyectos de acción colectiva (Moll, *et. al.* 2006: 15.626).

Estas adaptaciones, explicativas de los procesos de causación próxima del altruismo humano, invitan a pensar que han de explicarse por una causación última que conecta con la replicación del individuo, tanto en su aspecto genético como en lo que concierne a su propia identidad psicológica y cultural. La causalidad última de los comportamientos en animales adaptados para la vida social dentro de un nicho cultural es un puente que une la selección individual con la selección de grupo. Ese puente es la eficacia inclusiva poblacional, en este caso incluyendo no sólo replicadores genéticos,

sino también replicadores culturales determinantes del fenotipo orgánico del mismo modo que los estrictamente biológicos.

En el apartado 2 se exponen dos aspectos paradójicos de la cultura, que nos sitúan en la dimensión cultural del altruismo humano. Esta dimensión cultural del altruismo se explica por la capacidad que tienen los replicadores culturales de profundizar las semejanzas fenotípicas de los individuos, independientemente de su parentesco genético, según se explica en el apartado 3.

En el apartado 4 se analiza cómo se propagan los replicadores culturales a través del adoctrinamiento. El parentesco cultural explica la específica cooperación humana con fines culturales (apartado 4.1.), y la capacidad de participar en procesos de acción colectiva para la consecución de objetivos comunes de índole cultural. La especie humana rebasa las fronteras del coeficiente de parentesco genético en sus relaciones de cooperación, porque dispone de exadaptaciones apropiadas para activar los mismos procesos psíquicos, emocionales y cognitivos, en casos de relación con parientes genéticos y afines culturales (apartado 4.2.). La propagación de replicadores culturales es un factor explicativo de la cooperación de amplio espectro porque interviene en la generación de procesos de selección y evolución (apartado 4.3).

En el apartado 5 se aborda el instinto que hace posible la receptividad humana al adoctrinamiento cultural. Se trata del conformismo. Esta adaptación nos emplaza en la concepción del individuo no como fenotipo – vehículo único, que transporta replicadores exclusivos, sino como réplica de un agregado de rasgos culturales característicos de una identidad social, o sea, estamos dentro en la hipótesis de la selección cultural – grupal en la explicación del comportamiento altruista.

En el apartado 6 realizamos una aproximación básica a la definición y experimentación relativas a lo que podríamos llamar unidades básicas de replicación cultural, o memes.

Finalizamos el capítulo con la presentación de una teoría psicológica evolutiva del yo en el apartado 7, en la que se presenta la subjetividad como un proceso continuado de adscripción a replicadores culturales que conforman una identidad coherente del individuo. El sujeto proyecta su mismidad por identificación con determinados marcadores o estereotipos, y esta es la llave que abre la capacidad de ser altruista con todo aquello que se asume propio por proyección identificatoria, y por tanto constitutivo de incumbencia o “interés”. Paradójicamente, el interés propio del sujeto, entendido en un sentido identificatorio, es lo que explica su comportamiento altruista dentro del ámbito de su identidad.

5.2.- Dos paradojas de la cultura, y su relación con el altruismo humano

La especie *Homo sapiens* es la especie cultural y cooperativa por antonomasia. Se trata de una especie intrínsecamente cultural, y además cooperativa en sentido extenso por ser cultural. Como afirma Lévi-Strauss (2008[1949]: 240),

... es imposible referirse, sin incurrir en contradicción, a una fase de la evolución de la humanidad durante la cual ésta, aun en ausencia de toda organización social, no haya desarrollado formas de actividad que son parte integrante de la cultura.

El individuo humano es pues un organismo biológico que dispone de una psicología adaptada para la vida social y, dentro de esta, para una compleja regulación cultural de la misma. La cultura consiste básicamente en creencias y reglas que dan sentido y significado a la acción humana, y la hacen coherente con la de los demás individuos sin necesidad de costosos esfuerzos cognitivos. Dice John Bruhn que

La cultura da forma a nuestra apariencia externa; pero es menos obvio, aunque de mayor importancia, que la cultura también da forma a nuestras diferencias internas, tales como nuestras metas en la vida, nuestro sentido de la justicia, y nuestros sentimientos personales. La cultura es la miríada de influencias que conforman el cuerpo y la mente paso a paso en todas las etapas de la vida (...) La cultura es análoga a un paquete emocional que incluye valoraciones, sentimientos, tendencias de acción, expresiones, y procesos corporales estereotipados. Cuando un evento genera una respuesta emocional, todos estos componentes resultan activados debido a su interdependencia. Cada acontecimiento activa un paquete cultural de creencias, valores y normas que está grabado y reproducido por la mente, dando forma y significado a nuestro comportamiento. (Bruhn 2005: 243).

La cultura, como fenómeno, es una creación colectiva que surge de la interacción y difusión social. Los genes interaccionan con la cultura en procesos coevolutivos, influyéndose en sus configuraciones respectivas desde la unidad de análisis del grupo, o incluso de la especie, y no del individuo. Resulta aclaratoria la afirmación de Popper y Eccles (1984: 48) en el sentido de que «la evolución cultural no es sino la prolongación por otras vías de la genética». Asimismo, Clifford Geertz explica que:

No existe naturaleza humana como ente independiente de la cultura (...) nuestro sistema nervioso central fue formándose en gran medida en íntima interrelación con ella (...) Somos, en suma, animales incompletos o inacabados que, gracias a la cultura, alcanzamos nuestra plenitud. (citado en Hayek (1982[1976]: 267ⁿ²³).

Desde estos puntos de vista nos aproximamos a la cultura dentro de una perspectiva poblacional que contempla la propagación de replicadores, tanto genéticos como propiamente culturales:

... cualquier elemento de la cultura humana en que se piense tiene repercusiones, más o menos directas, en la constitución genética de la humanidad. Los principios éticos o religiosos, los sistemas de enseñanza, la medicina y la higiene, las costumbres y modas dietéticas, las leyes positivas sociales y económicas, y, en general, cualquier costumbre o ley, tienen consecuencias genéticas porque afectan diferencialmente la reproducción de los seres humanos. (Ayala 1994: 229-230).

El fenotipo humano es la unidad de análisis sobre la que operan las fuerzas de selección que propician la evolución genética. Es obvio que el fenotipo es el producto de la intervención de más replicadores que los existentes en el genotipo, o sea, hay más replicadores relevantes en la evolución humana que los estrictamente genéticos (Oyama 2000). Por otra parte, está constatada una superior semejanza genética entre los diferentes grupos que la de sus características culturales³⁰¹. Sin embargo, lo que cohesionan (y enfrenta) a los grupos humanos más allá del parentesco biológico es la cultura, o mejor dicho, las estrategias culturales de relación social. Tales estrategias constituyen las identidades y los compromisos con ellas, y la adscripción a tales

³⁰¹ Efectivamente, los individuos de la especie humana comparten entre el 99,7% y el 99,9% de la secuencia del genoma (Henderson 2010: 60-61). Sin embargo, existe una apreciable diversidad en los rasgos culturales de dos individuos seleccionados al azar. Ello se debe a la particularidad de los procesos históricos de etnogénesis y selección cultural, dado que en la especie humana «las relaciones sociales son intervenciones en el medio que sesgan las trayectorias evolutivas creando un medio selectivo para unas u otras características» (Broncano 2004: 189).

identidades es una de las principales dimensiones la prosocialidad humana. Este aspecto invita a plantear que tanto la ejecución como la transmisión de los comportamientos prosociales empíricos pueden resultar mejor explicadas por tipologías en la estructura y organización de los grupos, o sea, por aspectos de índole cultural - identitario, que por el parentesco genético (Boehm 1993). El éxito en la transmisión de las normas que regulan las relaciones intra y extra grupales contribuye a explicar el éxito de la supervivencia por selección grupal, esto es, la supervivencia diferencial del individuo dentro de órdenes culturales.

Los humanos son animales constructores de nichos cognitivos de naturaleza esencialmente cultural, y además de la herencia genética, la principal herencia en la especie humana es la cultura:

Heredamos de nuestros antepasados no solo bienes y productos diversos de los que derivamos riquezas materiales, sino también ideas y pensamientos, teorías y técnicas, a las que nuestra inteligencia debe su fecundidad. (Mises 2009[1949]: 214).

Una faceta de este patrimonio hereditario es el compendio de creencias irracionales y fantasiosas que, además de una estructura estable de apriorismos metafísicos, constituye el cimiento de toda cultura, y concretamente de sus mitologías y religiones. Estas ideas, y su simbología representacional, no forman parte del mundo racional, pero su capacidad de motivar la acción humana en procesos de acción colectiva resulta extraordinaria y, como explica Gazzaniga (1993: 249) refiriéndose al caso concreto de las creencias religiosas: «La mayoría de las personas religiosas estaría dispuesta a morir por esta parte mágica, no sólo para mantenerla intacta, sino también para difundirla».

Otra faceta de los sistemas de creencias de las diferentes culturas es que imponen a los individuos severas obligaciones para consigo mismos y para con los demás. Se trata de los sacrificios en favor de la comunidad, así como numerosos rituales de iniciación y demostración de compromiso, generalmente costosos en términos de tiempo, recursos materiales, salud, bienestar, etc., en definitiva, de eficacia reproductiva. Piénsese, por ejemplo, en las obligaciones relativas al celibato o la selección de pareja, abstinencia de ciertos alimentos, prohibición de realizar actividades productivas en determinados períodos del tiempo, renuncia a las propiedades

personales, o las obligaciones relacionadas con la caridad u otras contribuciones unilaterales a la comunidad. Estos aspectos de la cultura parecen contrasentidos y desafían las ideas corrientes acerca de la racionalidad del individuo humano en su doble vertiente: la racionalidad entendida como proceso lógico-empírico radicada en la realidad, el pensamiento rigurosamente ordenado y el método científico y, por otra parte, la racionalidad entendida como la optimización individual. En efecto, los compromisos culturales y sus demostraciones sociales imponen a los individuos costosas cargas que *a priori* causan un perjuicio evidente a corto plazo, puesto que la lealtad incondicional a las normas y creencias de estos sistemas ideológicos tiene un coste directo y cuantificable, que invitaría a la defección siguiendo la lógica de la elección racional y la de la supervivencia del -organismo- más apto. Desde el punto de vista evolutivo la paradoja puede plantearse con los interrogantes que señala Sosis (2005: 72) al respecto de este tipo de fenómenos culturales:

¿Por qué invierte nuestra especie tanto tiempo y energía en actividades dolorosas o, al menos, desagradables? (...) si la evolución nos ha especializado en allegar pronto los recursos que necesitamos, ¿a qué viene ese despilfarro de energía invertida en dichas prácticas?

Una solución posible a este problema científico consiste en indagar si la paradoja es tal, puesto que utilizando un argumento típicamente evolutivo, «las estructuras sociales desventajosas, del mismo modo que las estructuras morfológicas mal diseñadas, no pueden sobrevivir largo tiempo» (Gould 2010[1977]: 287). Desde este punto de vista, la interiorización de las normas de la cultura puede tener sentido desde una perspectiva de la racionalidad no centrada en el organismo-a-corto-plazo, sino en los replicadores-a-largo-plazo. Adicionalmente, las ideas fantasiosas y el pensamiento mágico pueden ser capacidades adaptadas que motivan conductas cooperativas relativamente más eficientes que la estrategia dominante, tanto para el individuo como para el grupo, y desde la perspectiva de la eficacia biológica, así como de una racionalidad ecológica:

Existen numerosas situaciones en las que puede ser adaptativo distorsionar la realidad. Incluso las creencias completamente ficticias pueden resultar adaptativas, dado que motivan comportamientos que son adaptativos en el mundo real. (Wilson 2002: 41).

Esta paradoja nos saca del patrón de racionalidad centrado en la optimización individual a corto plazo, y nos emplaza en un modelo de racionalidad ecológico – poblacional a largo plazo, centrado en la optimización de los resultados de la acción colectiva como ventaja adaptativa del individuo capaz de comprometerse con la (fantasiosa) ideología cultural. Así, Wilson propone que el principal desafío planteado por Calvino con sus ideas reformistas era unificar las numerosas facciones de la ciudad de Ginebra dentro de una unidad de propósito y de acción: su iglesia:

La iglesia de Calvino incluía un código de comportamientos adaptados al ambiente local, un sistema de creencias dentro de la mente del creyente que motivaba poderosamente para ejecutar ese código, y una organización social que coordinaba y reforzaba el código tanto para el caso de los líderes como de sus seguidores. (Wilson 2002: 111).

La cultura se erige pues como interacción constructora de sistemas sociales de destino común como bien común, y eso se consigue a través de la acción colectiva, en este caso la acción colectiva tutelada por la religión: «El calvinismo consiste en un sistema integrador con un propósito: unificar y coordinar a una población de individuos para lograr un conjunto común de metas a través de la acción colectiva» (Wilson 2002: 118). ¿Cuál es la medida del éxito de los sistemas integradores de destino común? Según Wilson, el calvinismo logró a través de la religión que la ciudad de Ginebra cobrara una importancia a nivel internacional ampliamente desproporcionada respecto a su tamaño, lo cual benefició sobremanera a sus ciudadanos a título individual. La homogeneidad cultural del calvinismo favoreció la unidad de propósito y la cooperación para su logro en la acción colectiva. La homogeneidad cultural es, según el autor, una precondition para que un grupo pueda actuar como unidad adaptativa, generando procesos de selección grupal, o sea, de acción unificada y coordinada de los individuos en orden a conseguir el éxito en la propagación de los replicadores con que se identifican. De este modo, lo que estaría propiciando un éxito diferencial de las normas de cooperación social interiorizadas en los cerebros es la supervivencia diferencial de los grupos, como sistemas culturales y no como agregados de individuos específicos y, dentro de ella, la de sus individuos entendidos como fenotipos homogéneos portadores de cultura. O sea, tenemos una teoría explicativa de la acción humana basada en la hipótesis de la selección de grupo.

Lo que explica la supervivencia y el éxito (económico y demográfico, por ejemplo) de los kibutzs israelíes de orientación religiosa respecto a los seculares (Sosis 2005: 74), el avance fulminante del cristianismo durante la Alta Edad Media en Europa (Ekelund y Tollison 2011), la relativa prosperidad de las comunidades amish contemporáneas en Filadelfia (Wesner 2010); o la relativa prosperidad de determinados grupos sociales en el siglo XX, entre los que figuran las élites educadas en el complejo “oxbridge” británico, el complejo “Army – Navy” o las redes sociales de Yale en EE.UU., las comunidades pacifistas e igualitarias cuáqueras o huteritas, u otros grupos profesionales de las clases medias occidentales (véase Akerlof 1983) es que, contraintuitivamente, en situaciones tipo del dilema del prisionero, u otros dilemas sociales, sus miembros tienden relativamente menos a elegir la estrategia dominante. De este modo, los comportamientos prosociales motivados por valores tales como la lealtad, el patriotismo, la honestidad, el coraje, o simplemente, disponer de unos principios claros y firmes de entrega personal a un determinado grupo o causa, favorecen a las personas que los poseen, a través de la inclusión económica y social en grupos donde los individuos comparten y valoran especialmente una identidad comprometida. ¿Por qué el compromiso personal con estos valores adquiridos culturalmente no se diluye frente a las fuerzas competitivas del oportunismo y el interés propio? Akerlof (1983: 58) lo explica con un argumento educativo, que es la replicación cultural y su repercusión en la selección de grupo:

... los padres de clase social acomodada enseñan a sus hijos a identificarse con otras personas acomodadas, y a ver a los menos afortunados como “distintos” (...) (de este modo) las personas no pueden pretender identificarse con otros miembros de su clase sin serles leales en la práctica. (Paréntesis añadido).

La hipótesis del altruismo selectivo y restringido a los que practican la norma de reciprocidad, así como la del éxito evolutivo de los individuos que se significan por una calidad diferencial (manifestada a través de conductas altruistas), ya fueron analizadas en el capítulo precedente, y se demostró que tienen capacidad explicativa de la evolución del altruismo desde una perspectiva de selección individual. En el capítulo actual se aborda el estudio del altruismo explicado por la interiorización psicológica de una identidad social, que impone sacrificios y restricciones al supuestamente pretendido comportamiento del individuo optimizador. La longevidad de los grupos como tales y

de las ideologías que los cohesionan, así como la supervivencia y el éxito económico de sus individuos, se explican por el nivel de exigencia de los sacrificios y restricciones impuestos por la cultura. Obviamente, tales muestras de altruismo se circunscriben, salvo excepciones tales como la caridad universal que predica el cristianismo, principalmente a las relaciones con los afines. De este modo, lo que aquí se plantea es la tesis de coevolución del comportamiento altruista por selección de grupo. Citando a Simon,

... si un grupo humano concreto tiene una cultura que le dota con una eficacia (*fitness*) superior en la competencia con otros grupos, entonces, en la medida que esa cultura no impone presiones selectivas auto-destructivas (*self-defeating*) en el grupo que la posee, puede proveer la base para una genuina “evolución social”; por ejemplo, para la supervivencia de este grupo a expensas de otros. Esto sólo puede darse, por supuesto, si es posible evitar la hibridación a gran escala entre los grupos competidores. La conquista europea de Norteamérica nos proporciona tal vez el ejemplo más claro de este proceso en los tiempos modernos. (Simon 1983: 57).

Admitido que los elevados estándares de compromiso exigidos por la identidad explican el nivel de cohesión intragrupal, *cæteris paribus* puede esperarse que los grupos más exigentes sean aquéllos más efectivos en generar altas tasas de cooperación interna. De modo que, como explica Irons (2001: 299), cuanto mayor sea la necesidad de cooperación, o la dificultad para generarla, mayores serán las exigencias de compromiso que establezcan las instituciones culturales³⁰². Si la especie humana dispone de la adaptación consistente en crear compromisos culturales y señales de reconocimiento de los mismos, que activan emociones motivadoras para cooperar con intensidad dentro de grupos identitarios, ¿cuáles son los procesos de causación próxima y última que permiten explicar esta adaptación, y qué capacidad explicativa tienen estos en la evolución del comportamiento altruista?

³⁰² Esta conclusión ilustra los esfuerzos de la innovación cultural e institucional por conseguir un ambiente social de sostenida (y sostenible) *tensión cooperativa*. A fin de evitar un deterioro de la sociedad con la proliferación de conflictos internos causados por comportamientos de optimización individual, los grupos pueden fijarse metas más ambiciosas que requieran un mayor nivel de compromiso de más personas durante más tiempo. La guerra y la conquista, las competiciones deportivas y la confrontación política, ideológica o científica, son estrategias orientadas a generar tensión cooperativa mediante proyectos de destino común, e incitar una cooperación intensa intragrupal. El caso paradigmático de la selección de grupo que más se ha estudiado es el conflicto bélico (véanse p.ej. Eibl-Eibesfeldt 1982, y Bowles 2006, y 2008), aunque la teoría puede aplicarse a cualquier otra actividad humana en la que se produzca una fuerte identificación psíquica del individuo con parámetros simbólico - culturales.

5.3.- Semejanza fenotípica: la hipótesis relacional y cultural de la cooperación humana

En el *Tratado de la Naturaleza Humana* David Hume realizó una observación acertada a los ojos de la psicología social contemporánea:

... la relación de sangre produce el lazo más fuerte de que el espíritu es capaz en el amor de los padres a los hijos, y en menor grado cuando la relación (de consanguinidad) disminuye. No sólo la consanguinidad tiene este efecto (...) Amamos (también) a nuestros compatriotas, a nuestros vecinos, a los que hacen un trabajo análogo al nuestro, a los que ejercen nuestra profesión y hasta aquellos que llevan nuestro nombre (...) Es manifiesto que los hombres se unen según sus peculiares temperamentos y disposiciones (...) Cuando notan la semejanza ésta actúa del mismo modo que una relación, produciendo un enlace de ideas. (Hume 1992[1739-1740]: II, ii, 4, pp. 228-229; paréntesis añadidos).

Si los lazos de sangre no son siempre los únicos factores explicativos del comportamiento prosocial motivado por la emoción del amor, como se reconoce actualmente, sino que también intervienen otros lazos fundamentados en la simpatía cultural, ello invita a estudiar la naturaleza y el contenido de la causalidad próxima que son tales sentimientos, así como su intervención en la motivación del comportamiento altruista.

Que la cooperación humana, especialmente en su modalidad de altruismo, se da profusamente en la familia genealógica, es un fenómeno empírico demostrado por la ciencia en los campos de la psicología evolutiva, psicología social, la antropología y la etología humana. Sin embargo, la teoría estándar de la eficacia inclusiva no es válida, de forma notoria en la especie humana, para explicar monocausalmente el altruismo porque el enfoque genético-inclusivo no alcanza todas las dimensiones del altruismo empírico humano. Los humanos no reconocen a sus parientes biológicos por rasgos identificatorios sensoriales³⁰³. Aunque es cierto que la herencia genética de los progenitores se expresa en rasgos morfológicos y de temperamento semejantes, sin embargo esto no siempre es así, se manifiesta tardíamente, o los rasgos mencionados no aportan suficiente información acerca de un emparentamiento inequívoco³⁰⁴. Dado que

³⁰³ Esto justifica que, como ilustra Richard Alexander (1971: 100), el interés de las personas y de los sistemas de regulación cultural por conocer las relaciones de parentesco, haya sido una «obsesión casi fanática».

³⁰⁴ La no existencia de feromonas u otros marcadores para el reconocimiento de parientes genéticos en la especie humana conduce a pensar que durante la mayor parte de la evolución humana esos marcadores no fueron necesarios para la eficacia inclusiva. Es muy probable que el reconocimiento visual, la memoria, y la habituación de proximidad en pequeños grupos semiaislados haya sido la manera de

los humanos infieren el coeficiente de parentesco de *proxies* fenotípicas, entre ellas las culturales, no es posible una cognición cien por cien fiable del grado de parentesco genético. Ello requeriría verificar con las modernas técnicas de estudio del ADN, no disponibles hasta fechas recientes en la historia de la humanidad.

La teoría genetista de la eficacia inclusiva ha sido criticada desde la antropología (un ejemplo es Sahlins 1976). Según esta perspectiva la cooperación en la especie humana está cimentada en la cultura y las relaciones culturales, se extiende mucho más allá del reducido círculo familiar, y esa flexibilidad permite establecer relaciones de cooperación fuerte incluso con personas desconocidas. La posibilidad misma de un altruismo limitado exclusivamente a las redes familiares se pone en entredicho en una especie en que la replicación sexual ha de ser exogámica, dados los conocidos efectos genéticos deletéreos de la endogamia³⁰⁵. Las bases de la cooperación, según la antropología, parten no tanto de las identidades genotípicas como de la experiencia de interacciones recurrentes en el tiempo, o sea, experiencias de proximidad relacional.

En efecto, en *Homo sapiens* la cooperación se extiende de la parentela genealógica a la relacional-comunitaria³⁰⁶. Aunque cabe reconocer la existencia de altruismo incluso anónimo y despersonalizado, sin embargo el tipo de relaciones reiteradas y estables son el punto de partida más frecuente para la generación inicial de identidades fenotípicas en los grupos humanos (Tyler y Blader 2003). Tal es el denominador común que puede explicar una cooperación étnico-cultural, y esto se trata, según los defensores de la teoría de la evolución cultural, del fundamento más importante de la cooperación social humana, o al menos un fundamento con tanto poder explicativo como el genético (Richerson y Boyd 2005)³⁰⁷. No obstante, y como ya se ha

admitir a los descendientes en el grupo, e integrar a toda la familia en bandas demográficamente estables y geográficamente móviles, que de manera esporádica llevaban a cabo intercambios reproductivos con otras bandas (véase Chapais 2010).

³⁰⁵ La reproducción en la especie humana ha de ser necesariamente exogámica a largo plazo, ya que una reproducción con parientes con quienes se comparte una elevada afinidad genética (p.ej. hermanos y hermanas, padres e hijas) concluye en defectos biológicos que dañan de forma irreversible la salud e impiden la replicación futura. Como consecuencia de ello la especie humana tiene una relativamente alta homogeneidad genética, que permitiría explicar la cooperación sobre otras bases científicas y sin necesidad de reducirla a las relaciones parentales.

³⁰⁶ Véase Richards (1987: especialmente Introducción y cap. 5).

³⁰⁷ Dentro de las ciencias sociales, el «no sólo por los genes» (*not by genes alone*) de Richerson y Boyd es la hipótesis por la que hasta ahora se han ido decantando en mayor medida los economistas que han planteado funciones de utilidad explicativas del altruismo familiar, y de la cooperación y el altruismo

explicado, el altruismo exogámico puede repercutir indirectamente en la eficacia genética, siempre que los sentimientos y emociones que lo motivan sean de tipo especular e inciten reciprocidad.

Las identidades fenotípicas no siempre se generan tras largos y complejos procesos ritualizados de aproximación. Existen dispositivos neuropsicológicos que permiten que una inclusión accidental del individuo en un determinado grupo, y en determinadas circunstancias, como por ejemplo, que se detecte otro grupo diferenciado en las inmediaciones, activa procesos de identificación psicológica intra grupal -basados primariamente, *cæteris paribus*, en la proximidad espacial- que motivan a cooperar³⁰⁸.

Los estudios de psicología social acerca de los procesos de etnogénesis coinciden en que los grupos étnicos y los sentimientos de pertenencia a grupos a través de identificaciones singulares, tienen su origen en sencillos procesos de aglutinación cooperativa de individuos cultural y geográficamente próximos ante la amenaza de, o la confrontación con, otro/s grupo/s diferenciados. Existe evidencia de que muchos de estos procesos de convergencia y cohesión grupal no dependen tanto de previos condicionamientos culturales o biológicos (al estilo de marcadores simbólicos o bioquímicos) como de interacciones reiteradas y el conocimiento mutuo. La convivencia, por ejemplo, tras un encuentro accidental (en una convalecencia hospitalaria, una oficina o un lugar de vacaciones), conduce frecuentemente a establecer relaciones sociales densas e identidades grupales que estimulan la cooperación y los mecanismos de selección grupal a través de diversos sesgos psicológicos de integración social, como la familiaridad, la amistad, la semejanza sobrevenida, o las circunstancias comunes (Hoffman 2000: 206 y ss.)

dentro de las identidades sociales (véase Becker 1997[1976a], y 1981; Bisin y Verdier 2001, y 2008; Akerlof y Kranton 2000, y 2005).

³⁰⁸ Esto puede informar de que, previo al desarrollo de las capacidades mentales más elevadas, como el raciocinio, la conciencia y el sentido moral, la especie humana se vio sometida durante el tiempo y las condiciones suficientes a fuertes presiones selectivas para la adaptación a cooperar dentro de un grupo no necesariamente familiar; que ese grupo, que pudo nutrirse de engrosar migraciones de individuos y fragmentos de otros grupos abandonados a su suerte, debió de contar con un tamaño ecológicamente óptimo; y que la pertenencia al grupo no fuera una cuestión sino de supervivencia por existir elementos de destino común que afectaran a todos sus miembros.

El famoso experimento de Muzafer Sherif y su equipo en el campamento de Robbers Cave resulta esclarecedor al respecto. En el verano de 1953 dos grupos de chicos (varones y todos con 11 años de edad) fueron llevados a sendos campamentos en un experimento controlado por el equipo de investigación de Sherif³⁰⁹. Ambos campamentos se ubicaron próximamente, si bien los adolescentes lo desconocían de antemano. Los chicos procedían de la misma ciudad, aunque no se conocían entre ellos previamente al campamento. Todos compartían rasgos culturales (protestantes, blancos, de clase media). En un momento dado se informó a los chicos de la existencia del otro campamento, y se les concedió autonomía para relacionarse. Las consecuencias surgieron de inmediato: los dos grupos se desafiaron mutuamente a enfrentarse en competiciones deportivas, ataques a las respectivas instalaciones y otras provocaciones que, no tardando mucho, desembocaron en una manifiesta enemistad entre los dos bandos y una situación de tensión que podría haber desembocado en la violencia si no hubieran intervenido los organizadores. En todo este proceso los dos grupos, autodenominados “Serpientes” y “Águilas”, desarrollaron una iconografía representacional con la que los miembros se identificaban apasionadamente: «Las banderas no sólo se capturaban, sino que también se quemaban y se hacían jirones» (Appiah 2007: 111). Esta dinámica no se puede considerar aislada; otros estudios etnográficos comparativos han proporcionado una notable cantidad de ejemplos ilustrativos de que las dinámicas de cooperación intragrupal se aceleran e intensifican en la medida que paralelamente lo hacen las de antagonismo intergrupal³¹⁰.

La clave que explica la evolución del comportamiento altruista se amplía pues a la selección de todos aquellos organismos de los que se puede extraer una dinámica relacional benéfica, y ello sin que la reciprocidad condicional sea el motivo *ex ante* de la acción: «ayudamos a las personas que disfrutan ayudándonos y ayudan a nuestros parientes» (Pinker 2004: 514). Esto conlleva aceptar que los marcadores de identidad no

³⁰⁹ El experimento fue publicado originalmente en 1961 por el Institute of Group Relations de la Universidad de Oklahoma. Aquí tomamos la información de Appiah (2007: 111-114).

³¹⁰ La cooperación y la solidaridad se incrementan en los grupos cuando se cree que existe una amenaza externa que es común, como un riesgo ecológico o la constatación de la amenaza de otro grupo rival. Esta dinámica cohesiva, que puede tener un trasfondo y una intencionalidad claramente económicas (el mutualismo, por ejemplo), y por tanto resultar eficientes desde la racionalidad utilitaria, sin embargo, a veces consiste en la expresión de tendencias instintivas profundas cuyo sentido económico instrumental no es tan evidente, o no lo es a corto plazo. La oposición reactiva a un grupo externo es un estímulo suficientemente motivador para cooperar.

están necesaria ni completamente predeterminados al nacer, sino que se asumen y van adquiriendo aquellos asociados a una supervivencia relacional. Asimismo, la asimilación de marcadores, y la estimulación emocional que estos producen, tienen una explicación epigenética y pueden funcionar como exadaptaciones de instintos implantados originariamente para la cohesión espacial de linajes familiares. De ello que:

Los individuos humanos están naturalmente predispuestos a cooperar con aquellos que se significan mediante señales fiables como semejantes (...) los individuos actúan ante sí mismos y ante los demás como ejemplares del grupo en lugar de como individuos diferenciados. De este modo, la identificación grupal y no el propio interés puede ser la base del comportamiento cooperativo (Caporael, *et. al.* 1989: 692-693).

La identificación con las personas y la identidad grupal explican los comportamientos altruistas humanos como conductas inclusivas exofamiliares (Dawes y Thaler 1988: 195; Loch, *et. al.* 2006: 224-225). La propia naturaleza de los procesos identificatorios, al resultar afectada por la semejanza entre sujeto observador y sujeto observado, asocia la conducta prosocial a coordenadas inclusivas de carácter al menos cultural, si no genético (a través de la ingeniería inversa); y ello aun asumiendo que las relaciones se producen entre individuos que no se consideran *a priori* genealógicamente emparentados³¹¹. La causa próxima es una estimulación ocasionada por un rasgo analógico, que puede venir o no explicado por una relación homológica³¹².

Los modelos explicativos de la relación altruista humana de nuestros días siguen haciendo hincapié en la explicación relacional – comunitaria y la identificación a través de rasgos de semejanza adquiridos en la enculturación (*in-group favoritism*). El proceso enculturador aporta cruzamientos multidireccionales de replicadores, tanto genéticos

³¹¹ Adviértase que el factor semejanza no es incompatible con la teoría genetista, como se pone de manifiesto en los estudios de cooperación entre individuos que comparten apellidos. En consonancia con el pronóstico de Hume, Oates y Wilson (2002) han encontrado pruebas de que datos de identidad tales como el nombre y/o apellidos funcionan empíricamente como señales fiables de semejanza. Compartir nombre y/o apellidos incita en una proporción significativa de personas (que alcanza el 30 por ciento, y más en el caso femenino que masculino) a realizar gestos de altruismo con individuos completamente desconocidos.

³¹² Como explica Gil-White (2001: 524), la relativa semejanza que motiva la identificación interpersonal no es, ni puede ser, la percepción de una estricta semejanza numérico-taxonómica. Puesto que la cantidad de rasgos identificables en un individuo humano es prácticamente ilimitada, la identificación por semejanza no puede estar respaldada sobre bases completamente empíricas. Más bien, se trata de una adaptación psicológica simplificadora de la realidad, que se basa en lo que los cerebros consideran semejante. O sea, «lo que importa no son los predicados compartidos, sino los *representados*» (*ibid.*; cursivas en el original). Lógicamente, los rasgos de semejanza son percibidos y clasificados por un órgano biológico adaptado, si bien los rasgos relevantes a considerar en cada momento histórico tienen un alto componente educativo y cultural.

(caso de emparejamientos sexuales) como culturales (caso de las influencias ideológicas y transferencias recíprocas de normas, costumbres y conocimiento). Estos cruzamientos (asimilables a la difusión por contagio, esponjamiento o deriva) activan patrones de acción cooperativa basados en el destino común del *pool* de replicadores compartido: el que da origen a la aproximación cooperativa, y el que se genera como consecuencia de una interacción cercana, íntima y prolongada. El factor semejanza se muestra decisivo previa y posteriormente a la relación:

Los parecidos entre los amigos, o entre dos individuos que tienen una relación social íntima (como marido y mujer) y han vivido juntos bastante tiempo, son en parte una expresión de lo que han aprendido el uno del otro, y en parte también de unas afinidades que pueden haber dado origen a su unión. (Cavalli-Sforza 2010[1996]: 184).

Esta constitución mixta (biológica y cultural) de la identidad aporta la comunidad de rasgos homogéneos y por ello se entra en una típica dinámica de destino común: los rasgos se propagan produciendo semejanzas, pero también las semejanzas se comparten para cooperar buscando objetivos comunes de manera más eficiente. Todo ello desencadena la activación de las causas próximas que impelen a cooperar con los semejantes, sin que la motivación sea premeditadamente de carácter utilitario:

Las teorías de la relación social a gran escala parecen identificar casi siempre ciertas relaciones íntimas o basadas en una “esfera comunitaria compartida” en la que las fronteras individuales se difuminan (...) La mayoría de la gente experimenta las amistades como relaciones comunales diferentes de las relaciones de intercambio, dotadas de una “estructura profunda” distinta de la “estructura superficial” del intercambio. (Nesse 2001: 31).

Tales aspectos guardan una relación estrecha con el factor confianza y su conexión con los innatismos relacionados con la preferencia por lo semejante. En efecto, en psicología social es bien conocido que las emociones que motivan para la inclusión social en grupos de identidad favorecen la cohesión por la vía de la percepción de una semejanza subjetivamente acentuada, así como expresiones de familiaridad y pertenencia tales como atribuir una superior confianza, honestidad, generosidad e interés recíproco a los miembros del grupo respecto a los *outsiders* (Hoffman 2000: 206 y ss.) Cualquier rasgo compartido, aunque se trate de un asunto trivial (como la marca de la ropa, el color del cabello o el acento lingüístico), resulta suficientemente activador, en determinadas circunstancias, de los sentimientos de favoritismo cooperativo hacia las personas que comparten tal semejanza, aunque sean individuos completamente desconocidos. Se trata

del fenómeno denominado “efecto del grupo mínimo” (*minimal group effect*). Estos procesos emotivos, que definen un “sesgo de familiaridad” o “favoritismo inclusivo”, contribuyen a reforzar la cohesión interna independientemente de si se considera o no la existencia de una amenaza externa común, en cuyo caso la cohesión y la identificación con el grupo se incrementan de manera exponencial (Hurlbert, *et. al.* 2000).

En otro orden de cosas, es conocido que cuando las personas están reiteradamente expuestas a un determinado estímulo, desarrollan e incrementan la tolerancia y preferencia por dicho estímulo (que puede consistir en un objeto inanimado, o en otro individuo *per se*) hasta el punto de crearse una relación de familiaridad e interdependencia³¹³. En las relaciones interpersonales humanas se produce este sesgo de familiaridad, y es lo que explica la superior identificación empática que se da con individuos con quienes se comparte una mayor proximidad. La cohesión interna dentro de los grupos de identidad equivale a una reducción de los estímulos a la competencia intragrupal. Una vez constituidos los grupos de rasgos homogéneos, e incluidos y aceptados en el grupo los individuos, las identidades grupales se refuerzan por la permanencia y cercanía en estructuras sociales estables, caracterizadas por la repetición de interrelaciones en el tiempo, el trato confiado e íntimo, y el consecuente personalismo relacional (Zajonc 1968; Kunst-Wilson y Zajonc 1980; Aknin, *et. al.* 2011). Existe evidencia experimental de que la interacción comunicativa explica la aparición de una identidad social básica, y asimismo que la comunicación en sí misma, no necesariamente directa y cara a cara, contribuye a reducir la distancia social entre individuos que son desconocidos entre ellos (Bochet, *et. al.* 2006; Bohnet y Frey 1999b). La evidencia experimental acredita la relación positiva entre las variables comunicación y cooperación (Balliet 2010).

Las estructuras sociales espontáneas, ya sea surgidas de la habituación relacional de proximidad, de preferencia por la semejanza o, con frecuencia, de la interacción de

³¹³ Este fenómeno es denominado en psicología “efecto exposición”, y consiste en que el incremento de la tolerancia, la confianza y la afectividad positiva resultante de la presentación reiterada de un estímulo anteriormente desconocido (Zajonc 1968). La receptividad al estímulo se incrementa cuando la exposición es inconsciente, favoreciendo procesos automáticos de habituación (Kunst-Wilson y Zajonc 1980). El efecto exposición se aplica a estímulos diversos (visuales, olfativos, gustativos, auditivos, etc.), y también se cumple con estímulos sociales, como exposición a estímulos y símbolos representativos de una identidad cultural. En estos casos se habla de desarrollo de preferencia por la familiaridad o sesgo de familiaridad (Hoffman 2000: 206).

ambos factores, activan procesos de pertenencia a, e identificación con, grupos y organizaciones formales que motivan para cooperar e incrementar la cooperación (Ashforth y Mael 1989). En este sentido, son un elemento explicativo de procesos de selección grupal, que se fortalece ante la existencia de objetivos comunes. Asimismo, la habituación y la preferencia por la semejanza tienden a aumentar la estabilidad, la confianza y la fidelidad de las relaciones interpersonales (McPherson, *et. al.* 2001: 436). Esto genera *preferencias de exclusividad* (valencia de hogar, en lenguaje de la etología; vinculaciones interpersonales, en psicología), así como la fidelidad a un conjunto de rasgos fenotípicos que caracteriza las relaciones estables dentro de la propia identidad personal y social.

Pasemos a explicar la causalidad última y las adaptaciones operativas que hacen posible la propagación de los rasgos culturales propiciatorios de la semejanza fenotípica. Una relevancia específica en este aspecto es la que tienen los procesos de adoctrinamiento cultural en las sociedades humanas, que existen porque son la expresión institucional epigenética del instinto humano de conformismo.

5.4.- Adoctrinamiento cultural para la convergencia fenotípica

Los grupos humanos comparten elementos diversos y variables de destino común, entre los que figura la supervivencia interdependiente de sus integrantes. El destino común que cohesionan los grupos puede ser un producto *a priori* de las propias preferencias innatas de la especie, que al ser homogéneas y de no existir rivalidad en su provisión, pueden ser susceptibles de satisfacción conjunta a través de la acción colectiva grupal. En otro orden de cosas, el destino común, teniendo una base epigenética *a priori*, puede ser construido como un *a posteriori* de la naturaleza humana, modelando las preferencias a través de la educación y el adoctrinamiento de los fenotipos. Esta “remodelación” cultural de los individuos aumenta la semejanza fenotípica, y concentra su destino común integrándolos en grupos culturales con objetivos compartidos, claros, coherentes, y delimitados. La acción instructiva del adoctrinamiento cultural es un factor impulsor de la cooperación a través de la convergencia fenotípica, dentro de la cual se incluye la interiorización de normas que

refuerzan el compromiso con una determinada identidad cultural. Esta fidelidad incluye tanto las categorías sociales en abstracto (los “modos de ser”), como los individuos afiliados a estas categorías (los que son “de un determinado” modo de ser).

Dado que la cooperación es una estrategia eficiente para conseguir los objetivos comunes, de ello se deriva el hipotético desarrollo evolutivo del vínculo a ciertos marcadores que señalizan una determinada identidad grupal. En el caso de la transmisión cultural se trata de marcadores simbólicos cuya existencia viene reforzada en la dinámica social, siempre que contribuyan al logro de metas comunes aportando ventajas diferenciales en relación con otras alternativas de organización social. En lo que se refiere a metas individuales no interdependientes, salvo que estén asociadas y resulten complementarias de otras metas (sociales) más amplias, los marcadores no aportan ventajas de supervivencia, y por ello no intervienen ni en la motivación ni en la señalización.

De los procesos de acción humana surgen dos fuerzas de selección enfrentadas, una en la línea de potenciar los requerimientos comunitarios sin elección racional, puesto que tales requerimientos fortalecen la eficacia de *todos* los individuos:

Si un grupo humano se mantiene delimitado por un propósito compartido o identidad común, entonces dispone de un conjunto de conocimientos que ordenan las relaciones entre los miembros del grupo. Ningún grupo es completamente autónomo; cada uno integra o es integrado en otros grupos y otras configuraciones de las relaciones humanas, y esas configuraciones tienen cada una de ellas su propia estructura o lógica. Usamos estas relaciones sociales estructuradas -esos solapamientos de conocimiento compartido- para coordinar nuestro comportamiento en el espacio y tiempo. De esto se sigue que la elección y la acción humanas están hasta cierto grado social o culturalmente condicionadas. (Ostrom y Ostrom 2003);

y otra en la línea de saltarse tales requerimientos y aprovechar cualquier oportunidad que confiera una ventaja diferencial al individuo:

Toda criatura social se halla en un equilibrio tenso entre lo que es sacar todos los beneficios y sufrir los costes que supone la vida en grupo. Esa situación crea una presión para obtener un balance correcto gracias al hecho de ser más listos. (Pinker 2004: 254).

El comportamiento oportunista es un factor generador de incertidumbre social, ya que

... los jugadores del ajedrez social tienen que ser más hábiles que los jugadores de este antiguo juego de mesa, porque las piezas no sólo cambian impredeciblemente de identidad -los caballos se convierten en alfiles, los peones en torres, y demás- sino que a veces los aliados cambian de color y se convierten en el enemigo. Los jugadores del ajedrez social deben estar constantemente alerta, al tanto de una ventaja potencial, vigilantes para detectar una desventaja inesperada. (Leakey 2001: 96).

Como la actitud de “pasarse de listo”, y la incertidumbre relacional que propicia, activan una carrera armamentista con un coste evidente para el individuo, y no sólo un posible beneficio, por ello en toda población que comparte metas comunes tenderán a evolucionar normas limitadoras del comportamiento bajo la forma de proyectos ordenados y regulados de destino común. En organismos que viven dentro de nichos cognitivos, como los humanos, tales proyectos se concretan en la cultura, que es un conjunto coherente y ordenado de normas y significados, y la uniformidad que esta genera en quienes la comparten porque la han asimilado (Nanay 2010: 84).

La uniformidad cultural se propaga mediante procesos variados de evolución cultural, uno de los cuales consiste en una vía de evolución de la cooperación que es el adoctrinamiento, voluntario e involuntario y, dentro de este último, se incluye igualmente el adoctrinamiento coercitivo. El adoctrinamiento es una transmisión e interiorización de cultura que se realiza básicamente a través de las capacidades innatas de aprendizaje e imitación. Los procesos sociales de convergencia cultural elevan el grado de semejanza poblacional, y permiten ampliar conocimientos a través de la imitación y, paralelamente, la acción de aprender e imitar expande el nivel de semejanza dentro de la población. El adoctrinamiento establece una homogeneidad fenotípica que posibilita la evolución de la cooperación fuera del enclave familiar o el deme de interacción densa que se trató en el capítulo 3.

La cooperación por selección cultural requiere la invención y difusión de marcadores culturales homogéneos que determinen tanto las metas de la población, como los procedimientos admitidos para conseguirlas. En entornos de incertidumbre éstos aportan señales de credibilidad sobre las expectativas atribuidas al marcador:

Los primates superiores que viven en grupos están en condiciones de reconocer individualmente a los miembros del grupo por sus características personales, y por tanto no

necesitan marcas grupales. El ser humano sólo las necesita a nivel secundario, cuando su comunidad crece más allá de la pequeña agrupación que es propia de su naturaleza hasta convertirse en una gran sociedad anónima. (Eibl-Eibesfeldt 1996: 96).

Efectivamente, la diferencia que hace singular a la especie humana es la riqueza y complejidad de sus marcadores de identidad grupal, un fenómeno de naturaleza completamente cultural. Tal diferencia posibilita la categorización de individuos en tipologías culturales y, si el marcador funciona eficientemente, ello posibilita anticipar un patrón de comportamiento endocultural, predecible y cooperativo (en la medida en que al menos se atiene a las normas de la cultura), en la interacción con desconocidos que señalizan de forma creíble la pertenencia grupal (Dunbar 1993: 732).

Los marcadores grupales exogámicos consisten fundamentalmente en un conjunto estructurado de creencias, normas y expectativas de comportamiento diferenciadas y específicas del grupo. Tales ingredientes a menudo son representados mediante marcadores perceptibles arbitrarios, como la gesticulación, vestimenta, temas de conversación y jerga utilizada, gustos estéticos y culinarios, modificaciones fisonómicas, etc., que permiten identificar fácilmente a los afines, y además con ciertas garantías de veracidad (Efferson, *et. al.* 2008: 1.844-1.845). Aunque el imaginario resulta discrecional, sin embargo el sustrato que hace de ello un aspecto relevante en la identificación de afines y la cooperación tiene un fundamento epigenético. Esto se explica porque la parafernalia cultural consiste en convencionalizaciones, extrapolaciones, o estandarizaciones -si bien a gran escala- de los patrones de relación predominantes en el primitivo demente de interacción densa (Tomasello, *et. al.* 2012: 685). Es algo así como una proyección cognitiva y emocional del clan primigenio. La explicación biológica reside en que existen fuerzas evolutivas propicias para que la cultura humana se expanda a través del adoctrinamiento, propiciando la homogeneidad en la población:

... los marcadores observables posibilitan al individuo realizar inferencias estadísticas sobre lo que, aun siendo inobservable, es lo verdaderamente importante; dígame las expectativas sobre el comportamiento de los demás en situaciones sociales de equilibrios múltiples. Cuando esto es así, el interaccionar preferentemente con aquellos que expresan un mismo marcador reduce la probabilidad de descoordinación e incrementa los resultados esperados. (Efferson, *et. al.* 2008: 1.845-1.846).

Un recorrido histórico hipotético pudo consistir en que tras una larga etapa evolutiva de enclaves dunbarianos, en los que la coordinación de pequeños números se asentaba en los factores interdependencia y visibilidad, la evitación de las trampas malthusianas a través de factores exógenos, como pudo ser la suavización de las condiciones climáticas, o endógenos, como las innovaciones materiales o institucionales, pudieron ocasionar expansiones súbitas de la población. Esto, unido al establecimiento de relaciones con otras poblaciones, hizo necesarias nuevas oleadas de innovación cultural tanto para el establecimiento de marcadores efectivos, propiciatorios de la cooperación, como para el reconocimiento de la fidelidad de los mismos. Los instintos originariamente adaptados para la cohesión de la pequeña banda hubieron de exadaptarse para la cohesión de agregados sociales de mucha mayor dimensión, sin lo cual el avance de la complejidad social no hubiera sido posible en la especie, como es el caso de las demás especies de primates. Pasemos a explicar estos fundamentos del grupalismo exogámico y exodémico característico de *Homo sapiens*.

5.4.1.- Marcadores de identidad: “Not by genes alone”³¹⁴

En la selección de parentesco, anterior a la selección cultural, la evolución puede haber adaptado ciertos instintos para detectar marcadores que señalizan una identidad biológica común. Es el caso, por ejemplo, de secreciones olorosas, emisión de determinados cantos o sonidos, exhibición de ciertos colores o el parecido físico. Estos marcadores invitan hipotéticamente a seguir la conducta “coopera con quienes se parezcan a ti en los atributos de olor, sonido, color, anatomía o fisonomía”. Otro marcador relacionado con la selección de parentesco puede ser el asociado a la proximidad, especialmente en grupos pequeños y sedentarios, aislados geográficamente. En este caso, el dictado del instinto es “coopera con quienes estén cerca de ti”. Como toda especie, la humana se vale de etiquetas o marcadores de señalización de identidades genotípicas, si bien con límites cognitivos que influyen en el alcance e intensidad de la cooperación.

Los humanos disponen de algunas habilidades cognitivas de reconocimiento relativamente desarrolladas, y aunque su olfato, oído o vista no disponen de la eficacia

³¹⁴ El título del subapartado es una alusión a la obra de Richerson y Boyd (2005), que tiene ese mismo título.

discriminadora que sí existe en otras especies, sin embargo se valen del parecido físico, la proximidad durante la juventud y el aprendizaje social de las normas y modelos de interacción social, que actúan como filtros señalizadores de la frontera que delimita a los que viven dentro del contorno de la parentela. No obstante, la disposición congénita de los animales sociales a seleccionar a favor de ciertos rasgos de afinidad, y no de otros, no se explica por una premeditada y abnegada selección inclusiva *a priori*, sino por una regulación más compleja que apunta hacia la cooperación con aquellos organismos que aportan eficacia desde la perspectiva del replicador. O sea, la selección de afinidad conlleva apearse (consciente o inconscientemente) a todo individuo que pueda contribuir de forma verosímil a mejorar la eficacia dentro de unas capacidades cognitivas evolucionadas, aunque de sí inherentemente limitadas³¹⁵.

En la especie humana la cooperación se produce en varios niveles de intensidad, frecuencia y altruismo, por lo que los sistemas de reconocimiento de identidad operan en varios niveles. La psicología del desarrollo prueba que los niños y niñas pequeños expresan una disposición congénita a la preferencia por individuos concretos, así como un fino sentido del reconocimiento de los individuos a los que conocen y tienen apego o sienten una previa atracción (Hood y Bloom 2008). La expresión de preferencias discriminatorias es un rasgo natural de la especie. Existen variados sistemas de reconocimiento grupal, dada la importancia de la interacción estable dentro de diversas estructuras grupales que caracteriza a los primates. La evolución ha seleccionado un modelo de comportamiento humano basado en la cooperación dentro de grupos diversos con diferentes grados de estabilidad y frecuencia, reciprocidad y generosidad, confianza y compromiso. Cada esfera de cooperación tiene sus propios marcadores de identidad, algunos de los cuales son genéticos, como el parecido físico, si bien la mayoría no son predeterminados, sino aprendidos y transmitidos culturalmente:

Es imprescindible tener en cuenta que las pruebas indican que estos marcadores no vienen determinados genéticamente, al menos no como se suele interpretar. Lo que está claramente codificado en nuestros genes es la tendencia a detectar determinados marcadores y la capacidad para aprender a hacerlo, pero no la forma específica que éstos adquieren en una cultura determinada. Nuestro cerebro está preparado para captar determinadas señales, pero las aprendemos una vez hemos nacido. (Navarro 2004: 14).

³¹⁵ En los primates no se han encontrado adaptaciones para el reconocimiento de parientes. Como subrayan Castro Nogueira, *et. al.* (2003: 121), el reconocimiento de la prole se lleva a cabo mediante un criterio de proximidad, no de señalización.

En efecto, algunas especies han desarrollado capacidades para señalar la identidad y/o la compatibilidad genética a través de la emisión y captación de feromonas específicas (caso de diversas especies de roedores, por ejemplo). Como ocurre en el caso de los demás primates, hasta hoy no se ha reconocido en la especie humana una adaptación especializada en el reconocimiento de parientes. No obstante, existen investigaciones recientes que avalan la existencia de una capacidad genética, compartida con otras especies de mamíferos, de reconocer, a través de los olores corporales, aquellos individuos con quienes se tiene una superior histocompatibilidad³¹⁶. Esta adaptación no está orientada directamente al altruismo a través de la selección de parentesco, aunque sí indirectamente, además de proveer de información necesaria para cooperar con los individuos más compatibles en la tarea de la replicación genética. Lógicamente, esto supone una ventaja adaptativa para seleccionar de manera eficiente las parejas sexuales con que se replicará con mayor éxito parte del material genético.

En otro orden de cosas, los estudios sobre antropología sensorial permiten avanzar en la línea de la replicación cultural, desde la constatación de que en la especie humana se da una atracción discriminatoria hacia los individuos que tienen olores corporales similares a los del receptor (Howes 2003). Esto, que no se apartaría de la interpretación puramente genetista, abre camino a la consideración de la cultura en la selección y emparejamiento cooperativo entre individuos, dado que los olores corporales resultan explicados no sólo por la genética, sino además y principalmente por la alimentación y hábitos de vida. Tal constatación tiene una destacable dimensión étnica y cultural, y plantear la hipótesis de una asociación de replicadores culturales y biológicos en la selección de individuos con quienes cooperar. Dicha selección se guiaría por las variables de compatibilidad y semejanza.

En efecto, la especie humana se vale más que ninguna otra de marcadores puramente culturales conforme se incrementa la complejidad de los órdenes sociales:

³¹⁶ La investigación de referencia es la de Wedekind, *et. al.* (1995). Recientemente se ha publicado el trabajo de Davis (2013), quien profundiza en esta teoría. McNulty, *et. al.* (2013) han demostrado en un estudio longitudinal que las actitudes, valoraciones y sentimientos inconscientes revelados implícitamente por los recién casados, resultan más predictivas y explicativas del éxito o fracaso de la relación conyugal que las justificaciones y supuestas actitudes deliberadas.

... la cultura implica marcadores para la identidad grupal que resultan especialmente relevantes en la definición de lo que forma y no forma parte del grupo (...) El etnocentrismo puede aportar una barrera de separación efectiva para la difusión de elementos culturales dentro del nivel de la cultura. La clase, casta, género, ocupación e incluso los grupos de entretenimiento están simbólicamente demarcados dentro de las sociedades. Dentro de los límites del grupo, aun siendo un grupo grande, los matrimonios, la difusión de conocimientos, y otros procesos mixtos contribuyen a generar uniformidad cultural dentro del grupo, así como importantes diferencias entre grupos. (Boyd, *et. al.* 1997: 366).

La lengua, y dentro de la lengua el acento y el dialecto, son claramente marcadores fenotípicos de identidad cultural. Es un hecho constatado la cooperación supeditada a marcadores de identidad, de manera que la perfección de marcadores comunes está asociada a una cooperación potencial de mayor nivel e intensidad. En el estudio de Cohen (2012) se analiza la influencia del acento, dentro de una lengua, como marcador eficaz de identidad, relevante para explicar una cooperación intragrupal diferencial³¹⁷. Ejemplos de marcadores de identidad cultural no lingüísticos que pueden citarse son las prácticas culturales de circuncisión o mutilación genital instauradas en algunas religiones, que son señales de compromiso altamente costosas y generalmente impuestas por la presión social. Prácticas de carácter voluntario son los tatuajes, cicatrices simbólicas, estilos de vestimenta u otros ornamentos, como el peinado o las deformaciones anatómicas, además de aprendizajes ritualizados de ciertas habilidades, creencias, sistemas ideológicos, etc., como el caso de la memorización de mitologías y ceremonias. Todos los grupos humanos, desde las pandillas de adolescentes hasta los ejércitos y congregaciones religiosas, clubes deportivos, partidos políticos o asociaciones científicas, comparten un capital social de prácticas simbólicas que permiten el reconocimiento entre sus integrantes, y la señalización de pertenencia a una misma tribu extendida.

Algunas características definitorias de un marcador de identidad fiable, ya sea cultural o biológico, son: perceptible, asociado al individuo, comparable, difícilmente falsificable, costoso, discriminable, dinámico, ancestral, adquirido tempranamente, y universalizable (Cohen 2012: 594-600). Los buenos marcadores culturales de identidad comparten con los marcadores biológicos muchos de estos atributos, que se muestran

³¹⁷ Téngase en cuenta que esto no se contradice con la teoría sociobiológica de la eficacia inclusiva. Esto se debe a que el acento es un marcador difícilmente falsificable, que se adquiere mediante la prolongada exposición a una lengua materna. Además se trata de un marcador que informa de la procedencia étnica y geográfica, que puede resultar explicativo de un coeficiente de parentesco (genético y cultural) superior al promedio en interacciones aleatorias de poblaciones complejas.

eficaces en la discriminación de falsos afines infiltrados. Esto hace que ya no sólo la presencia de los marcadores, sino su cuantía, intensidad y exigencia en la observancia, sean aspectos relevantes desde el punto de vista de la señalización de veracidad y compromiso. Así, tal y como explica Irons (2001), un atributo fundamental en la garantía de veracidad es la necesidad de una fuerte inversión previa de eficacia en la adquisición del marcador, así como la entrega (figurada o real) de fianzas y otras formas de expresión de compromiso auténtico, obediencia y lealtad al compromiso con una determinada identidad cultural. Irons (2001: 298) lo expresa como «onerosos rituales de iniciación». Es obvia la procedencia filogenética de la base de innatismos que soportan estas prácticas, características de una economía de los compromisos creíbles, como se refleja en la siguiente descripción que hace Ruse (1983: 64) de los procesos de inversión parental en activos específicos que se dan en numerosas especies:

En la estrategia de la felicidad doméstica, la hembra obliga al macho para que haga una inversión sustancial anterior a la cópula. En el momento que la copulación tiene lugar, el macho está fuertemente comprometido con la hembra, pues ha tenido que construir un nido, alimentar copiosamente a la hembra, realizar un laborioso rito nupcial y otras cosas parecidas. Básicamente, ahora el macho no podría interesarse en abandonar, pues la siguiente hembra con la que se encontrase le exigiría nuevamente ese esfuerzo.

Otra modalidad de compromiso reforzado, en este caso con un elevado coste de oportunidad por la ruptura del compromiso, es la estigmatización de los que abandonan o son expulsados del grupo, tanto de parte de los ex-compañeros, como de los potenciales acogedores. Explica Urrutia Elejalde (2014: 55) que quien se aparta de su grupo suscita desconfianza en otros grupos, dado que ha señalado su capacidad de disidencia y de «traicionar» una identidad:

... no basta con salirse de un grupo para pasar a un limbo desidentificado, sino que, no hay más remedio que pasar de un grupo identitario a otro, y a esto se le llama traición, entendiendo esta como la disidencia *in acto*.

Efectivamente, el poderoso sentido de pertenencia al grupo, y los remordimientos de conciencia que ocasiona la idea de “traicionar a los suyos”, tienen una explicación evolutiva en que la fidelidad es un rasgo adaptativo ya que cualquiera que abandona un grupo es objeto del resentimiento y de la venganza de sus ex-compañeros, y dado que necesita ingresar en otro, sufrirá la desconfianza aplicada al que se sabe que es capaz de

disentir y traicionar (Urrutia Elejalde 2014: 53). Consecuentemente, los costes de la disidencia son un factor de estabilidad y orden social, máxime cuando las conductas de fidelización refuerzan el compromiso asumido con el grupo a través de inversiones en marcadores de carácter indeleble. Muchas de estas inversiones son altamente costosas, carentes de valor fuera del grupo o negativamente valoradas, y frecuentemente son irrecuperables e incluso inocultables debido a la imposibilidad de un “borrado” *a posteriori* de los marcadores. El celibato irreversible, por ejemplo, exigido como prueba de lealtad religiosa o profesional, señala no sólo una elevada capacidad de compromiso por el coste implícito que se asume de no replicación biológica, sino además por la carencia de lazos familiares en que cobijarse ante una situación circunstancial de aislamiento social por abandono o expulsión del grupo. Los tatuajes, incisiones en la piel, deformaciones óseas, etc., delatan al traidor ante los enemigos del grupo que abandona. Asimismo, los conocimientos adquiridos, rituales aprendidos y valores interiorizados delatan una determinada visión del mundo que, en ocasiones, puede tener valor fuera de su previo entorno social, pero que también en ocasiones, y dependiendo de los costes impuestos por la ortodoxia previa y futura, pueden llegar a suponer un lastre o un estigma. La interiorización, permanencia y difusión de los marcadores culturales en los vehículos orgánicos tiene pues una explicación evolutiva que pasamos a comentar, empezando por los mecanismos activadores de la cooperación dentro de un *pool* de replicadores culturales.

5.4.2.- Exadaptación y señuelos culturales

Gintis, *et. al.* (2005: 26) han presentado indicios que permiten apoyar la tesis de una pronta cooperación exogámica y exodémica de tipo epigenético. En efecto, la cooperación con desconocidos no sólo puede haber evolucionado por razones reproductivas obvias en los grupos humanos (evitar los efectos adversos de la endogamia genética en los grupos humanos), sino que además debió de ocurrir en tiempos muy remotos en la historia de la especie, y por adaptaciones grupales y protoculturales. La justificación se encuentra en que, como explican los autores, se ha demostrado que las características etnográficas, circunstancias sociales y condiciones ambientales del Paleolítico superior determinaron crisis demográficas de tipo catastrófico en intervalos de alrededor de tres décadas. Por lo tanto, la longevidad y

estabilidad de los pequeños grupos de cazadores-recolectores debió ser muy mermada por la inestabilidad de las condiciones de supervivencia³¹⁸. Una consecuencia destacada de esto, *cæteris paribus*, es una esperanza de vida menor: según los experimentos paleoantropológicos realizados, se situó en torno a los 17 años. Estas precarias condiciones de supervivencia necesariamente hubieron de provocar desmantelamientos apresurados de los grupos y frecuentes migraciones de los supervivientes a otros grupos diferentes, o reagrupaciones de individuos dispersos pertenecientes a unidades diezmadas. El constante flujo migratorio de individuos sólo pudo ser posible si los grupos receptores acogían a los extraños y cooperaban con ellos, lo cual supone necesariamente la coevolución de capacidades emotivo-cognitivas de tipo inclusivo hacia los desconocidos, que debieron requerir refuerzos culturales capaces de producir un efecto señuelo³¹⁹. Por *efecto señuelo* entendemos la capacidad de la cultura para estimular el surgimiento de experiencias emotivo – cognitivas del tipo de las relacionadas con la interacción de la *kin selection*, pero fuera del enclave de interacción densa. Obviamente se trata de una exadaptación explicativa de la faceta más característica de la sociabilidad humana, y este “engaño” a los hipotéticos genes del parentesco y la proximidad ha demostrado tener una fuerte capacidad adaptativa³²⁰.

La cultura favorece los procesos de convergencia fenotípica, y ésta incrementa el agregado poblacional emplazable dentro de la que podríamos denominar “frontera de posibilidades de cooperación”. Como explica Mises:

³¹⁸ Comenta Gat (2010: 200) que las condiciones existentes durante la larga etapa evolutiva de la humanidad, durante la cual la forma de vida fue la trashumancia de pequeñas bandas dunbarianas de cazadores y recolectores, hacían que «el espectro del hambre y (muerte por) la inanición estuviera siempre presente» (paréntesis añadido).

³¹⁹ Esto matiza la tesis defendida entre otros por Richard Alexander (1975) sobre la inherente hostilidad intertribal, o sea, el ataque sistemático a los desconocidos como rasgo evolucionado al servicio de la cooperación intra-tribal. El principal problema que tiene la hipótesis del conflicto inter-tribal consiste en explicar por qué evolucionaron los instintos de hostilidad hacia los desconocidos si la cooperación con ellos, por ejemplo a través del comercio de bienes y el intercambio de técnicas, destrezas y conocimientos, puede aportar ventajas adaptativas superiores a las que aporta la agresión y la guerra (Richerson y Boyd 2008). En otro orden de cosas, tal y como teorizó Kropotkin en *La ayuda mutua*, los animales no solo cooperan para competir contra otros grupos o para regular conflictos intragrupales, sino que también, y destacadamente en el caso humano, lo hacen para sobrevivir en ambientes naturales hostiles, modificando el medio e influyendo impremeditadamente en su propia evolución.

³²⁰ Pedersen (2004) explica que el mismo sistema neuroendocrino que activa las emociones adaptadas para el cuidado de la prole y la vinculación con los parientes consanguíneos, permitió a los primeros humanos sentir afectos prosociales, que producen vínculos y comportamientos cooperativos con los miembros de otras familias de la tribu, y, llegado el caso, con individuos desconocidos.

Los miembros pertenecientes a una nación o rama lingüística, o los grupos que dentro de ella se forman, no están siempre unidos por sentimientos de amistad y buena voluntad (...) Fue ideológico el impulso que permitió superar los antagonismos, inspirando a los miembros de la nación o grupo lingüístico aquellos sentimientos de comunidad y pertenencia que los actuales nacionalistas consideran fenómeno natural y originario. (Mises 2009[1949]: 201).

La función que desempeñan la ideología, el credo organizativo, banderas, lemas y escudos, himnos, uniformes, y la identidad grupal en un sentido extenso, en la cohesión grupal y el aumento de la cooperación dentro de agregados poblacionales amplios ha sido destacado desde diversos campos de la ciencia social (North 1984, y 2005; Denzau y North 1994; Simon 2001; Richerson y Boyd 2001). Así, que existan grupos diferenciados no implica la agresión y la guerra como *modus vivendi*, ni tampoco la inexistencia de replicadores transversales propiciadores de convergencia fenotípica. Estos replicadores, como las actuales creencias en los derechos humanos o la demostración de la hermandad genética universal de la humanidad, son explicativos de conductas de tolerancia y apertura cooperativa³²¹. La percepción de un individuo perteneciente a otra cultura no implica necesariamente su hostigamiento o exclusión, puesto que las creencias y conocimientos culturales del tipo de los anteriormente citados pueden activar los mismos procesos psicológicos de empatía y confluencia mental que los propios de una relación parental (basada, igualmente, y en gran medida, en la creencias como la de consanguinidad, u otras como un lazo espiritual común). La cultura participa en la generación de lo que Tomasello, *et. al.* (2012: 674) denominan “mentalidad grupal”, de la que emana la posibilidad de que los individuos compartan una «intencionalidad para crear y mantener el complejo de tecnologías, sistemas simbólicos, e instituciones culturales de las sociedades humanas». La interiorización de prácticas culturales a gran escala produce el afloramiento de sistemas institucionales, los cuales facilitan (aunque no garantizan) que los grandes grupos operen de forma similar a los pequeños enclaves dunbarianos. Este metafórico “haplodiploidismo cultural”

³²¹ Y a la inversa, principios y creencias de segregación y rechazo ocasionan comportamientos de agresivos y excluyentes. En efecto, como ilustra Gil-White (2001: 518), los mismos individuos pertenecientes a la etnia judía que en la etapa histórica de la República de Weimar estuvieron plenamente integrados en la sociedad y la cultura germanas (en lengua, costumbres, vestimenta, etc.), y que incluso lucharon dentro del ejército alemán en la Primera Guerra Mundial, fueron considerados enemigos de la identidad alemana por la ideología antisemita del partido nazi. La persecución y ejecución del pueblo judío no sólo se justificó por una supuesta incompatibilidad de ambas identidades -que en numerosos casos no existía por haber renunciado numerosos judíos a sus tradiciones y costumbres-, sino que se fundamentó en la creencia esencialista, completamente ideológica, de una supuesta naturaleza corrupta de la estirpe judía.

humano, o hipotética eusocialidad cultural, se apoya en la creación cultural e interiorización de convenciones, normas e instituciones al nivel del grupo como unidad ontológica, y

... requiere una nueva manera de pensar en la que existe un “nosotros”, que constituye no solamente los socios actuales en una determinada empresa colaborativa, sino la integridad de todos los que forman parte de la sociedad. (Tomasello, *et. al.* 2012: 685).

Tradicionalmente el “humanitarismo” (la tolerancia cooperativa en el trato con los demás) ha sido considerado uno de los rasgos definitorios de la naturaleza humana, y ha contribuido a trazar la línea demarcatoria de los conceptos de civilización y salvajismo. Si bien la actitud humanitaria es filogenética, resulta reforzada por creaciones culturales de tipo ideológico, ceremonial y organizativo, que estimulan la biología activando las disposiciones inclusivas propias de la familia y el demente primitivo (Gintis, *et. al.* 2003: 165-166). Esto implica que es posible producir un “efecto señuelo” con la materia prima de los instintos adaptados a los ámbitos de proximidad relacional, para inducir una expansión de las emociones y sentimientos de pertenencia hacia agregados poblacionales de mayor tamaño. El dispositivo puede funcionar como sigue:

... las emociones son un fenómeno puramente biológico, en el sentido de que son producto de la acción de neurotransmisores. Se trata de imperativos biológicos tales como el hambre, el miedo o el deseo sexual (...) Al mismo tiempo, los ambientes sociales institucionalizados también generan y manipulan las emociones, con frecuencia utilizando rituales de dramatización que tienen efectos claramente directos y fácilmente comprensibles sobre las bases biológicas de las emociones. (Thornhill, *et. al.* 1997: 205)³²².

Esta es una particularidad característica del más destacado animal cultural, y aunque el comportamiento “civilizado” puede ser una construcción en alto grado cultural, no por ello deja de ser una consecuencia del efecto señuelo, asociada a aprendizajes y rituales culturales que excitan la producción por el sistema endocrino de los mismos neurotransmisores que produce la familiaridad consanguínea (Carter 2008). Como

³²² Las emociones tienen un sustrato genético de especie, pero en el caso humano su etiquetación y formas ritualizadas de expresión dependen de la cultura. Se trata por lo tanto de una construcción social, igual que ocurre con la expresión simbólica (lingüística o artística) de los sentimientos o de los conceptos. Bruhn (2005: 245) se refiere a las emociones como representaciones internas de normas sociales. Así, un mismo *qualie* emocional, como el amor o la tristeza, puede ser enculturado de maneras diversas: el silencio representa duelo en Japón, y ese mismo duelo se expresa con sonoro llanto en Irán. Del mismo modo, una expresión externa de emociones como es la risa, puede representar diferentes *qualia* emocionales en distintas culturas: respuesta de gratitud a un cumplido recibido en Vietnam, insinceridad en Rusia.

concluyen Tomasello y Vaish (2013: 251), las influencias culturales no crean por sí mismas las tendencias básicas de tipo prosocial y cooperativo que se aprecian en la primera infancia, sino que las moldean y adaptan a un determinado entorno social y cultural. De este modo, citando a Damasio:

La historia de nuestra civilización es, hasta cierto punto, la historia del esfuerzo persuasivo por extender los mejores “sentimientos morales” a círculos cada vez más amplios de humanidad, más allá de las restricciones de los grupos internos, y que eventualmente abarquen toda la humanidad. (Damasio 2005: 159).

Pasemos a reflexionar sobre el protagonismo desempeñado por la evolución cultural en la convergencia fenotípica.

5.4.3.- Naturaleza evolutiva de los marcadores culturales de identidad

La cooperación humana, y de manera especial su intensidad, distan mucho de ser aleatorias e indiscriminadas. Existe una correlación demostrada entre el sedentarismo relacional y la tasa de cooperación (Nowak. *et. al.* 2000: 18). Esto concuerda con que la homogeneidad asociada a la difusión de replicadores (genéticos y culturales) es una cuestión de carácter relacional y de proximidad. Sin embargo, un factor que limita la capacidad explicativa de la cooperación basada en el enclave dunbariano es la propia naturaleza móvil de las especies del reino animal. Muchas especies sociales son móviles, y además habitan ecosistemas sociales con poblaciones muy cuantiosas. La movilidad unida al factor cuantía incrementan la cantidad y complejidad estructural de las posibles interacciones, y con ello la complejidad social del ecosistema³²³.

El tamaño y la complejidad de la organización social, sin embargo, confieren ventajas de supervivencia a los individuos, como son las economías de escala y la especialización, la organización y combinación de factores, la capacidad de protección frente a amenazas y de provisión de recursos asociadas a órdenes sociales complejos (Gifford 1999: 134). De este modo, existen fuerzas selectivas que operan a favor del

³²³ Considérese que la cantidad de posibles relaciones sociales diádicas que puede generar una población sin restricción alguna viene descrita por la ecuación $R = n \cdot (n-1)/2$, donde n es el tamaño poblacional. Así, en una población de 150 individuos se pueden dar 11.175 posibles encuentros. La probabilidad de volver a repetirse un encuentro es $8 \cdot 10^{-9}$. En una población de 1.000 individuos la cantidad de encuentros posibles se incrementa de forma no lineal hasta la cifra de 450.000, siendo la probabilidad de repetir el encuentro $4,93 \cdot 10^{-12}$.

tamaño y la complejidad poblacionales³²⁴. No obstante, la incertidumbre asociada a la identidad (cooperativa o no) de los potenciales interlocutores aumenta la dispersión de la eficacia ganada, y puede reducir la eficacia media extraída de las interacciones. *Cæteris paribus*, la movilidad y el mayor tamaño poblacional, que son adaptaciones con valor de supervivencia, tienen la contrapartida de reducir la tasa marginal de cooperación aunque el resultado de la cooperación media pueda ser más eficiente³²⁵.

La manera de compatibilizar las ventajas y desventajas de habitar entornos caracterizados por la incertidumbre relacional ha consistido en adaptaciones psicológicas y de comportamiento adicionales, tales como el desarrollo de la autoconciencia, la absorción de normas culturales y la cognición interpersonal, la desconfianza inicial hacia los desconocidos, el interés por supervisar a los demás y conocer su pasado, la práctica del cotilleo y la atracción por las circunstancias de la vida privada ajena, etc. (Enquist y Leimar 1993; Yamagishi, *et. al.* 2005). Asimismo, las adaptaciones eficaces para comprender la estructura de los incentivos (que incluye el carácter de las intenciones ajenas), fraguar alianzas, recordar situaciones e identidades, procesar información social, etc., asociadas a una inteligencia social más evolucionada, se impusieron en la carrera armamentista intraespecífica como adaptaciones favorables a la evolución de la cooperación en situaciones de elevada incertidumbre ambiental (Dunbar 1996). Estas adaptaciones posibilitan la detección de potenciales contrarios y afines, amigos y enemigos, leales y oportunistas, y asimismo permiten señalar eficazmente el carácter, reconocer potenciales aliados e informar sobre marcadores de identidad interiorizados. Una inteligencia social de mayor nivel capacita para la receptividad al adoctrinamiento cultural. Ello se debe a que el adoctrinamiento interviene en el comportamiento a través de la homogeneidad cultural de los individuos y su influencia en la generación de efectos señuelo, así como en la generación de expectativas verosímiles sobre las intenciones y el comportamiento ajenos.

³²⁴ La historia humana acredita la superación evolutiva de las estructuras sociales menos complejas por otras más complejas, no así la generalidad de la historia natural (McNeill y McNeill 2010: 363).

³²⁵ El tema ha sido ya tratado en el apartado 3 del capítulo 3 de esta investigación.

Existe evidencia experimental de que en la especie humana la cooperación es una variable dependiente de, entre otras, las expectativas sobre intenciones y comportamiento de los demás. Cuando las personas creen que las motivaciones egoístas son universales, o simplemente que abundan en la población, tienden a reducir o suprimir su cooperación. Por el contrario, en ambientes sociales donde hay más cooperación los individuos cooperan con una mayor probabilidad (Traulsen, *et. al.* 2010). Esto es así porque el carácter de las creencias interiorizadas influye en la conformación de las preferencias y la disposición a cooperar (Batson, *et. al.* 1987). De este modo, los elementos reguladores y fenómenos de la interacción social que influyan en la conformación de expectativas y creencias pro-cooperativas sobre el entorno social, actuarán favoreciendo la motivación a cooperar. De ello que el marco institucional (*institutional framing*), en el que intervienen normas formales, ideologías, prejuicios populares, transparencia informativa, y nivel de tolerancia social a las conductas antisociales, juega un papel determinante en la configuración de las preferencias individuales. Éste explicaría económicamente la motivación a cooperar y los resultados de la cooperación³²⁶.

Dado que la cooperación es una estrategia provechosa y con valor de supervivencia, existen indicios teóricos explicativos de adaptaciones orientadas a la constitución de estructuras sociales de cooperación viable. Es el caso de las agregaciones de individuos culturalmente homogéneos, con o sin lazos de consanguinidad (Levine y Kurzban 2006). En estas dinámicas relacionales existen procesos de cualificación e identificación, que conforman redes de individuos identificados como afines en algún sentido relevante que permita inferir disposición a cooperar. La cualificación e identificación previas se llevan a cabo mediante procesos enculturadores que no dependen de la voluntad del individuo³²⁷.

En la interacción social entre individuos enculturados de forma semejante se producen sinergias y externalidades positivas, y los humanos se hallan psicológicamente

³²⁶ Véanse al respecto Frey (2006), y Ok y Vega-Redondo (2001).

³²⁷ Téngase en cuenta que la enculturación no es necesariamente un proceso voluntario, como la imitación de aquellas conductas que *parecen* desembocar en un mayor éxito para el logro de objetivos, sino que más bien consiste en una absorción inconsciente de conocimiento tácito, e incluso en alguna medida de una acción coercitiva no contractual como es la presión del entorno social (Boyd y Richerson 1993; Irons 1991).

capacitados para aprovechar beneficios relacionales a través de la búsqueda e identificación de los socios más propicios (*network externalities*). La cooperación y el altruismo dentro de estas redes viene a ser una consecuencia endógena derivada de la propia composición de la red. De ello, aparecen dinámicas de destino común (formales o informales) que afectan a los integrantes de dichas redes, caracterizadas por una difusión inherente de beneficios en cascada dentro de una red expansiva:

... en el ambiente de un actor focal existen otros actores que tienen objetivos que, por alguna razón, se alinean con los del primero. Si otro actor, el agente A, tiene objetivos similares a los del actor focal, éste registrará beneficios derivados de un bienestar continuado del agente A. Es importante destacar que tales beneficios *no* se derivan de la reciprocidad, sino que se trata de un subproducto de la persecución por parte del agente A de metas que se alinean con las del agente focal. Crucial en el modelo es que esta situación activa una dinámica en la que puesto que el agente focal tiene interés por el bienestar del agente A, surge un interés recíproco de A por el bienestar del otro. De este sentido de “compromiso profundo” derivado de la estructura del modelo, se deriva la particularidad de que la arquitectura cognitiva resultante asociada a las preferencias de relación interpersonal tenderá a incluir mecanismos diseñados para valorar a los individuos que, por cualquier razón, también nos aprecian a nosotros. (Levine y Kurzban 2006: 175).

Se trata pues de que las particularidades de los individuos contribuyen a construir estructuras relacionales fortalecidas por externalidades e intereses recíprocos. Esto es así porque los individuos extraen eficacia de la existencia de otros individuos en algún aspecto homogéneos, así como de las propias estructuras sociales (Tomasello, *et. al.* 2012: 680). De este modo, puede explicarse evolutivamente tanto el interés por el grupo y la comunidad como la búsqueda de individuos semejantes, la construcción cultural de sistemas que incrementan la afinidad, y la aparición espontánea de preferencias de simpatía hacia los afines, tales como la mentalidad afiliativa y la vinculación con el entramado de individuos, símbolos, creencias y significados definitorios de “lo propio” y “lo común”. Simon (1983: 9) deriva de estos entramados de intereses interdependientes y simpatías recíprocas una identidad colectiva o sentido de comunidad. Tomasello, *et. al.* (2012) aluden a una mentalidad grupal cohesiva. Otros autores, como Denzau y North (1994) y Greif (1994) se refieren al concepto de “modelos mentales compartidos”, entendidos como construcción social y patrimonio común. Estos modelos, que son construidos en la interacción social,

... sirven para construir, coordinar, y contextualizar la acción social, así como interpretar, planificar y recordar. La palabra “modelo” puede sugerir que dichos procedimientos de

coordinación relacional son primariamente cognitivos, pero no obstante integran además emociones, motivaciones, necesidades, juicios y actitudes evaluativas (...) las personas los utilizan irreflexivamente y sin ser capaces de articularlos con claridad. (Fiske 1998: 7).

En todos estos casos se alude a los elementos culturales, que transmitidos a través de procedimientos diversos contribuyen a fijar la cooperación como rasgo conductual dentro de la población. Las instituciones culturales humanas pueden de este modo incrementar o reducir comparativamente la dotación de recursos necesarios para la cooperación (que pasa a ser entendida como variable flujo, y no variable *stock*), y ese andamiaje institucional viene a ser un activo comunitario cuya acumulación ha pasado (y ha de pasar) la criba de los filtros de la selección natural y cultural.

Un destacable resultado producto de la transmisión cultural es que estar adoctrinado en una determinada mentalidad, o imaginario cultural, se asocia a un funcionamiento de la cognición y de la motivación sociales que se apartan del patrón de optimización según las reglas de la elección racional. Doguglass C. North explica cómo influye la ideología en el comportamiento humano, precisamente desviándolo de la estrategia de optimización individualista. Esto ocurre a través de una «percepción distorsionada» que surge como consecuencia de la comprensión y la valoración de los acontecimientos sociales del mundo, percibidos de acuerdo a unas coordenadas filtradas por la ideología aprendida:

La prima que un individuo establece sobre su coste de oportunidad antes de comprometerse en una acción ilegal es una medida del valor que deposita en la legitimidad (una consideración ideológica). Asimismo, el coste neto en que un individuo incurre al forzar el cambio social es una medida de la injusticia y alienación que él mismo siente. Se necesita algo más que un cálculo sobre el coste/beneficio individual privado para explicar el cambio y la estabilidad. Los individuos pueden ignorar ese cálculo en un intento de cambiar la estructura, debido a su profunda convicción ideológica de que el sistema es injusto (...) El cambio y la estabilidad en la historia exigen una teoría de la ideología para explicar las desviaciones producidas respecto al cálculo racional e individual. (North 1984: 26; paréntesis en el original).

En el plano de análisis no sólo cognitivo, sino que además contempla emociones y sentimientos morales, está confirmado que los estereotipos de señalización y reconocimiento, y la ideología que los categoriza y da significado, activan emociones en el sistema límbico que motivan a la aproximación confiada o, por el contrario, al rechazo-huída. Una señal de aproximación es una señal de familiaridad que inhibe el sistema del estrés y genera confianza y atracción. La identidad cultural y la

identificación, así pues, son posibles y motivan para el comportamiento progrupal porque las emociones humanas están adaptadas para activar emociones inclusivas, generar confianza y vincular a los individuos dotados de marcadores que expresan afinidad (Cloutier, *et. al.* 2011).

Como ya se ha explicado, los marcadores de identidad cultural pueden llegar a tener un coste y una complejidad elevados una vez han evolucionado para evitar falsificaciones, así como claras y sencillas especificaciones cuando el *pool* de señalizadores de la identidad cultural es amplio, heterogéneo y despersonalizado. No obstante, existen marcadores que se imprimen en el individuo inconscientemente y de manera cuasi indeleble, y que informan de manera inequívoca sobre sus orígenes sociales y culturales³²⁸. Ejemplos de estas características son la lengua, el dialecto y el acento, marcadores poco o nada costosos y fáciles de adquirir espontáneamente, que sin embargo pueden llegar a incluir señales sumamente complejas e idiosincrásicas sobre la trayectoria vital que resultan extraordinariamente difíciles de falsificar:

Después de la pubertad el producto de la transmisión cultural lingüística ya casi ha cristalizado, por lo menos en el aspecto fonológico. Cada individuo tiene su propio acento, que reproduce con pequeñas variantes el del ambiente geográfico y social en que ha crecido. (Cavalli-Sforza 2010[1996]: 198).

Los rasgos culturales tales como matices léxicos y fonéticos, costumbres higiénicas o alimentarias, actitud hacia lo novedoso, o hacia la práctica de doctrinas religiosas, filosóficas e ideologías en general, constituyen marcadores fiables de la identidad de los grupos a los que pertenece, o perteneció, el individuo. De esos grupos el individuo ha recibido sus genes y también su cultura, y por ello determinan tanto la afiliación sentimental como la identidad. Asimismo, y en lo que aquí concierne, de los grupos el individuo hereda modelos de comportamiento coherentes con su identidad, que permiten predecir la respuesta conductual de la persona ante situaciones culturalmente estereotipadas (Boyd y Richerson 1987: 66). Ahondemos en el instinto que hace posible

³²⁸ Expone Edgar Morin:

El *imprinting* fija lo prescrito y lo prohibido, lo santificado y lo maldito, implanta creencias, ideas, doctrinas que disponen de la fuerza imperativa de la verdad o de la evidencia. Enraíza en el interior de las mentes sus paradigmas, principios iniciales que mandan los esquemas y modelos explicativos, la utilización de la lógica, y ordenan las teorías, pensamientos, discursos. El *imprinting* va acompañado de una normalización que hace callar toda duda o contestación de las normas, verdades y tabúes. (Morin 2004: 304).

la difusión de replicadores culturales, y la cooperación endocultural, a través de los procesos de adoctrinamiento: el instinto de conformismo.

5.5.- Conformismo y replicación cultural

Todo animal cultural dispone de un cerebro adaptado para asimilar cultura, sin la cual este órgano no se desarrolla adecuadamente durante los denominados “períodos sensibles” de la ontogénesis³²⁹. El cerebro humano desarrolla conciencia a través de su ontogénesis relacional y social. La hipótesis de un cerebro en estado de naturaleza es una ficción, y la realidad empírica de tal conjetura son aberraciones del tipo niño de Aveiron o niños cruelmente apartados de la sociedad. Estos casos demuestran que

... los obstáculos y los cambios en el contorno social y cultural generan modificaciones de la estructura neuronal (...) los niños que crecen en estas situaciones ven profunda y en ocasiones permanentemente afectadas sus facultades cognitivas, muestran señales de retraso mental y carecen de habilidades lingüísticas. (Bartra 2006: 48).

En la especie humana este gregarismo, o conformismo social-cultural (*docility*), consiste en la tendencia innata a aprender de otros o, genéricamente, a exponerse e interiorizar de forma espontánea la influencia social (Simon 1990; Boyd y Richerson 1993). La adopción de la influencia social es un fenómeno psíquico en el que interviene la interacción de grupos de neuronas especializadas denominadas neuronas espejo (Rizzolatti, y Sinigaglia 2006; Iacoboni 2009). Los animales con capacidad de aprendizaje social tienen una marcada disposición al conformismo, que en el caso humano se lleva a cabo a través de procesos, formales e informales, basados en el mimetismo y sincronización gestual, fonética, sintáctica y emocional, entre otras (Wiltermuth y Heath 2009; Tanner, *et. al.* 2007). El conformismo es un rasgo filogenético que aporta ventajas de supervivencia al individuo, relativas a suplir las limitadas capacidades epistemológicas de la mente humana. Además, como en las demás especies de mamíferos sociales, el conformismo forma parte de la evolución de dispositivos neurológicos que capacitan para conductas de cooperación dentro de un orden social coherente y predecible. De este modo, el individuo conformista no optimiza de forma sistemática, sino que sigue normas grupales dentro de equilibrios

³²⁹ En estos períodos de la vida del individuo el cerebro se muestra especialmente receptivo, a través de su plasticidad, a la influencia del medio social.

conductuales de cooperación intraespecífica. Este coste de oportunidad soportado en términos de eficacia resulta compensado, entre otros aspectos benéficos, por el conocimiento social adquirido en los procesos de aprendizaje, así como la economía de recursos que conlleva la propia limitación social de la carrera armamentista intraespecífica. La herencia dual (genética y cultural), de este modo, es evolutivamente eficiente desde la optimización de los escasos recursos cognoscitivos y materiales³³⁰.

Los análisis contemporáneos de la capacidad humana para replicar cultura admiten que el instinto de conformismo es un instrumento al servicio de la replicación cultural, que aporta ventajas evolutivas en términos biológicos (Castro y Toro 2004). Esto se debe a que un *Homo sapiens* sin enculturar dispone de unas capacidades de supervivencia relativamente reducidas en comparación con un individuo enculturado³³¹. Por otra parte, un individuo sin arraigo normativo-cultural se mueve en la indeterminación de valores y conductas, y tiende a experimentar anomia consecuencia, entre otras causas, del conflicto psíquico que producen la falta de referencia moral, debilidad de la voluntad, disonancia cognitiva y la carencia de un sentido existencial de pertenencia (Gergen 1992). Estos desequilibrios están asociados a una falta de identidad personal madura, inestabilidad, e identidad fragmentada:

... en las reglas dadas que asignan a los hombres su lugar en el orden social y con él su identidad, queda prescrito lo que deben y lo que se les debe, y cómo han de ser tratado y contemplados si fallan, y cómo tratar y contemplar a los demás si los demás fallan. Sin tal lugar en el orden social, un hombre no sólo sería incapaz de recibir reconocimiento y respuesta de los demás; no sólo los demás no sabrían, sino que él mismo no sabría quién es. (MacIntyre 2004[1984]: 158).

La asimilación conformista de cultura no supone el abnegado sacrificio altruista del individuo por su grupo, ni requiere tampoco la existencia permanente de conflicto intergrupal. La capacidad para interiorizar cultura puede explicarse dentro de un modelo que recoge «un sentido débil de la selección de grupo», en el que el individuo recibe acriticamente -y no siempre de forma voluntaria o consciente- *imprintings*, o sea,

³³⁰ Tal economía de recursos cognoscitivos ha configurado una naturaleza humana dotada de patrones de acción fijos y aprendidos, que permiten tratar el comportamiento humano como una integración de procesos automáticos, controlados e interiorizados (Schneider y Shiffrin 1977).

³³¹ Los “niños salvajes” estudiados en los últimos siglos tienen una inferior esperanza de vida y posibilidades de replicación en su estado presocial. Asimismo, transcurrido un tiempo de estar apartados de la influencia cultural expresan grandes dificultades o incapacidad de asimilar cultura, lo que resta eficacia respecto al individuo enculturado. Véase Cavalli-Sforza (2010[1996]: 190 y ss.)

programaciones estereotipadas de preferencias (Gintis 2006a: 22)³³². Los buenos hábitos culturales (que han superado procesos de selección), los aprendizajes de técnicas y destrezas, y de reglas de coordinación estables, favorecen la eficacia del individuo que internaliza las normas de su entorno social³³³. No existe garantía alguna de que las normas culturales contribuyan a maximizar la eficacia desde la optimalidad global, si bien tienden a resultar eficaces desde la perspectiva de la optimización local:

Los agentes humanos disponen de funciones objetivo *socialmente programables*. El comportamiento humano no sólo depende de las *creencias*, lo que ya refiere a restricciones a la libertad de acción, sino además a *valores*, que son las verdaderas *metas* de la acción (...) Se suele suponer en los modelos evolutivos que los agentes actúan para maximizar su eficacia, sin embargo la perspectiva de la racionalidad limitada parte de que los individuos no maximizan *eficacia*, sino una *función objetivo* que ha evolucionado para reflejar, con mayor o menor exactitud, la eficacia biológica en un ambiente determinado. (Gintis 2003: 407-408; cursivas en el original).

Otra perspectiva para comprender la evolución cultural es la que podemos llamar, continuando con la terminología de Gintis, sentido fuerte de la selección de grupo la cual, admitiendo procesos simultáneos de selección multinivel, se basa en restricciones de orden superior a la selección individual³³⁴. En este enfoque los replicadores culturales no se difunden exclusivamente vía imitación impremeditada, sino que la asimilación de los mismos puede deberse a un aprendizaje coercitivo, como coercitiva puede ser también su observancia vía regulación social y castigo al disidente (Irons 1991).

En efecto, dado que el cumplimiento de normas reguladoras de una población es una señal de pertenencia a esa población, en la medida que la observancia de dichas

³³² El requisito de existencia de procesos de aprendizaje social, o algún tipo de socialización restringida al nivel del grupo, son aspectos explicativos de la homogeneidad intragrupal a la vez que relevantes para la activación de procesos de selección grupal. Aquí estamos considerando las fuerzas simultáneas de evolución asociadas tanto a la eficacia diferencial derivada de las cualidades del fenotipo individual dentro de su grupo, como la eficacia diferencial del sistema cultural que regula el comportamiento individual del grupo frente a los sistemas culturales de otros grupos (Sober y Wilson 2000).

³³³ El individuo humano no es un utilitarista calculador, sino que está adaptado para asimilar por aprendizaje cultural aquellas pautas de supervivencia que le permiten vivir en grupos adaptados al entorno y que, en interacción con otros grupos distintos, han sido capaces de desarrollar una superior capacidad expansiva tanto biológica como cultural (Hayek 1982[1976]).

³³⁴ Un nivel superior de selección grupal, o sentido fuerte de la misma, incluiría además otras fuerzas coevolutivas de selección, tales como el adoctrinamiento activo, el ostracismo de los que han realizado defección en el pasado, existencia de restricciones a la movilidad intergrupala, y una dispersión limitada de los individuos portadores de una herencia genética y cultural común. Véase Bowles y Gintis (2008: apartado 6).

reglas constituya un factor ventajoso de cooperación para el grupo, entonces el comportamiento tipo que no cumpla las reglas será castigado. Obviamente, el castigo al disidente a través de la presión social y la coerción tiene efectos tanto culturales (exclusión de la confianza cultural depositada en el individuo, e incapacidad de éste para replicar sus propios replicadores culturales) como directamente biológicos (reducción directa de su eficacia reproductiva genética). El conformismo en este sentido fuerte, o beligerante, es una forma de denominar el proceso de acatamiento y refuerzo del sistema de reglas, que incluye no sólo la coacción, sino incluso la violencia.

Los beneficios que obtiene el individuo procedentes de la selección en varios niveles, asociada a la pertenencia a un sistema cultural, no consisten sólo en los bienes individualizables que pueda proveer la cultura, como puede ser una herramienta o procedimiento técnico aplicable a la resolución de un problema concreto del individuo. La principal dotación de bienes que aporta una cultura a los individuos es la de los bienes públicos y comunes, como son el acceso comunal a recursos naturales e información o los hábitos saludables que, por ejemplo, evitan la propagación de enfermedades contagiosas, o la propia utilidad general que supone la existencia de una lengua común. La lengua permite al individuo obtener información y conocimiento que utilizará como bien privado, pero la lengua en sí misma es un activo del que no se puede apropiarse ni beneficiar ningún individuo en exclusiva. Las costumbres saludables y el uso de una lengua tienen un efecto directo en la replicación biológica. Tal y como apostillan en este aspecto Ekelund y Tollison (2011: 76), existe una hipotética «métrica darwiniana operativa» que explica el éxito de un conglomerado de bienes culturales como es el caso del cristianismo en sus primeros cuatro siglos de expansión cultural. Sostienen los autores que la cultura cristiana pudo inducir a comportamientos mejor adaptados a las circunstancias ambientales y sociales que los de otras religiones y organizaciones políticas. Así, mencionan la superior fortaleza de las redes de ayuda mutua endo y exofamiliares de los cristianos ante las adversidades y las plagas, así como la salubridad de la prohibición del adulterio, la cual redujo la transmisión de enfermedades y ensalzó a la mujer dentro de la familia cristiana como factor social de seguridad y estabilidad. Esto contribuyó no sólo a una superior cohesión familiar de las familias cristianas, sino también a una superior cohesión social de los grupos cristianos, que practicaban tanto la caridad como las restricciones morales a la promiscuidad

sexual. Consecuencia de ello, *los cristianos* (como población culturalmente diferenciada, y no como agregado de individuos independientes) se beneficiaron de una esperanza de vida superior a la de los paganos, tal y como evidencian los datos hallados en las inscripciones mortuorias:

... los frutos de esta fe no limitaban a los dominios del espíritu. El cristianismo ofrecía también mucho a la carne (...) La pertenencia era cara, pero de hecho suponía una ganga. Como la Iglesia pedía mucho a sus miembros, poseía recursos para *dar* mucho. Por ejemplo, como se esperaba que los cristianos ayudaran a los desposeídos, muchos de ellos recibieron tal ayuda, y todos podían sentirse seguros ante los malos tiempos. Como se esperaba que cuidaran a los enfermos y moribundos, muchos de ellos recibieron también similares atenciones. Como se les pidió que amaran a los otros, fueron amados en retribución. Y como se les exigía observar un código moral mucho más restrictivo que el de los paganos, los cristianos -especialmente las mujeres- disfrutaron de una vida familiar más segura. (Stark 2001: 204; cursivas en el original).

Otro aspecto cultural benéfico para la replicación biológica, y claramente favorable para la replicación cultural es, en este caso, el proselitismo de la fe, una acción difusora de replicadores que se vale de la imitación (sentido débil de la selección grupal), aunque también, como es sabido, de un adoctrinamiento impuesto y coercitivo (sentido fuerte de la selección grupal). De este modo, la superior capacidad de inclusión e integración de *outsiders*, característica del universalismo doctrinal cristiano, fue decisiva para alimentar un crecimiento cultural sin fronteras geográficas, políticas o étnicas, algo que no ocurrió, o no lo hizo en la misma medida, en otras religiones y filosofías de la época:

(El cristianismo) creció porque los cristianos se constituyeron en una comunidad intensa, capaz de generar esa “invencible obstinación” que tanto ofendía a Plinio el Joven (...) Y los medios esenciales de su crecimiento fueron los esfuerzos mancomunados y motivados del creciente número de creyentes cristianos, que invitaban a sus amigos, parientes y vecinos a compartir la “buena nueva” [...] el amor y la caridad se deben extender incluso mucho más allá de la comunidad cristiana (...) A mi juicio, la forma más importante mediante la cual el cristianismo sirvió como movimiento de revitalización dentro del imperio fue ofrecer una cultura coherente y *enteramente despojada del componente étnico*. Todos eran bienvenidos, sin necesidad de algún vínculo étnico. (Stark 2001: 225-226, y 231-232; cursivas en el original).

Los movimientos ideológicos y religiosos que han sobrevivido culturalmente y se han propagado entre poblaciones mayores son aquellos que consiguen implantar creencias y aprendizajes funcionales relacionados con el cuidado de la salud, la selección de pareja, la solidaridad con los afines y la seguridad relacional con los *outsiders* (Lester 2002). La utilidad de la religión y la política, y por extensión de todo

fenómeno cultural delimitador, y aglutinante y homogeneizador de creencias y comportamientos, consiste en su capacidad de generar conductas coherentes con la conformación y el logro de metas de destino común. Asimismo, su capacidad de establecer sistemas de control social eficientes en la organización de la acción colectiva dirigida a dichos objetivos. Los sistemas culturales existen porque posibilitan que los individuos logren unidos lo que les resulta imposible conseguir aisladamente. Incluso un individuo capacitado de forma excepcional para propagar sus replicadores biológicos y culturales tendría menos éxito que otro, peor capacitado, pero que opera dentro de un grupo cultural organizado.

Desde este enfoque el comportamiento altruista se define, además de como la que puede ser una norma social replicable como cualquier otra (p.ej. acudir a determinadas ceremonias y contribuir al sostenimiento de su parafernalia), también, y principalmente, como una restricción cultural epigenética, admitida por su valor intrínseco, a la consecución de óptimos individuales (no matar, no robar, no mentir, etc.). Dado que esta preferencia por la sociabilidad (la débil y la fuerte) aporta en sí misma ventajas adaptativas, no es necesario que exista un replicador biológico que codifique un supuesto rasgo altruismo ya que, admitida la evolución del conformismo, el altruismo es un subproducto evolutivo que viaja como un “autoestopista” de la previa adaptación prosocial y pro cultural de internalizar normas (Gintis 2003)³³⁵.

La sociabilidad adaptada de la especie humana tiene su soporte filogenético en la “absorción” de las identidades grupales por el individuo, principalmente las identidades locales más próximas y en los años de maduración del cerebro infantil. La ontogénesis de la asimilación cultural se especifica en dispositivos biológicos tales como cerebros, hormonas, extremidades, ojos, oídos y gargantas, y se vale de ellos como medios para replicarse (Richerson y Boyd 2001a: 175). Existen dos instancias en cuanto al soporte de los replicadores culturales. Por un lado están los instrumentos portadores de representaciones simbólicas en el proceso de comunicación replicadora entre individuos (planos, diseños, libros, modelos matemáticos, partituras musicales, etc.) Tales

³³⁵ El sistema de refuerzo filogenético consiste en que los conflictos producidos entre el cumplimiento de las normas morales y los deseos personales (comportamiento moral *versus* comportamiento hedonista) activan automatismos emocionales en el sistema límbico que repercuten, en individuos sanos, en un aumento de la ansiedad y el malestar psíquico (Sommer, *et. al.* 2010; Greene 2005).

instrumentos son vehículos exosomáticos en sí mismos susceptibles de replicarse con la intervención humana (p.ej. a través de la copia)³³⁶.

El replicador de la información cultural se almacena dentro de los organismos, o sea, se trata de los replicadores endosomáticos que son el soporte orgánico de la información dentro de la unidad biológica de transporte que es el organismo. Previo a esto se requieren codificaciones genéticas que hacen posible el soporte que a su vez soporta la replicación cultural. Este puede ser el caso, por ejemplo, del hipotético patrón cognitivo universal, de carácter precultural, de acuerdo al cual el 90 por ciento de los niños aún no enculturados creen en su propia existencia inmaterial anterior a la concepción y nacimiento (Emmons y Kelemen 2014). Sobre este patrón cognitivo se asientan las creencias culturalmente estructuradas en torno a la espiritualidad, y las ideas sobre una vida eterna incorpórea y lo divino. Las religiones, entre otras modalidades de sistemas de creencias culturales no asentados en los principios de la lógica empírica, son pues expresiones culturales epigenéticas cuyos replicadores culturales se asientan en cerebros biológicamente predisuestos (*ibid.*) Del mismo modo ocurre con la disposición congénita a aprender una lengua y el posterior aprendizaje cultural de la misma (Pinker 1995). Así, sólo el individuo que dispone de un cerebro educado para entender una determinada expresión lingüística, ya sea el lenguaje matemático, el chino mandarín, o el musical, puede llegar a “interiorizar” un teorema, un texto confuciano, o una sinfonía de Beethoven. Para quien no dispone de esa adaptación orgánico-cultural los replicadores culturales exosomáticos no existen o carecen de significado. Un ejemplo adicional es el de la educación emocional, aspecto sobre el que ya cabe poca duda de que aun siendo las emociones humanas filogenéticamente universales, y por lo tanto determinantes en buena medida del contenido de las normas sociales, sin embargo, también su expresión, reconocimiento y significado vienen determinados por el condicionamiento cultural (Elster 2001: 98 y ss.).

³³⁶ Según Bartra (2006: 185) estos vehículos exosomáticos son memorias artificiales que la humanidad ha construido socialmente a modo de prótesis que almacenan la memoria colectiva humana. Esto se explica por la comparativamente limitada capacidad de almacenamiento de información de que tienen los cerebros individuales.

Todo esto conlleva que el organismo ha de ser ontogenéticamente adaptado, o capacitado para ser un eficaz portador y transmisor de cultura. Como el proceso de adaptación puede ser costoso, laborioso y largo, se trata de una de las principales inversiones de alta especificidad de los individuos, y por la propia naturaleza de la adaptación se requieren:

- a) Estructuras biológicas portadoras de la cultura, como pueden ser las neuronas y la bioquímica específica de las conexiones sinápticas;
- b) innatismos adaptados para no ofrecer resistencia a la concreción de tales estructuras y a la entrada de información aferente (caso del instinto del lenguaje, las intuiciones acerca de lo espiritual o el propio instinto de conformismo), e
- c) innatismos adaptados para la vinculación y la lealtad a los grupos proveedores de la educación e información cultural interiorizadas por el individuo.

Este último aspecto, el de la lealtad a los grupos culturales es, según Simon, la base explicativa más importante del comportamiento altruista humano. La cultura -y todo el entramado biosocial de soporte y difusión- viene a ser un caso principal de generación de proyectos humanos de destino común que aglutinan a los individuos a través de la cooperación. Las causas próximas se refieren a la identificación e identidad psicológica de los individuos, y a todo el entramado sentimental cohesionante que se genera en torno a ello:

Las lealtades grupales no solo tienen un componente motivacional sino también cognitivo. Definen las fronteras del grupo representativo que se incluye dentro de lo que es considerado el bien común, y causan variaciones particulares y modelos explicativos compartidos del mundo que configuran el pensamiento de los miembros del grupo (...) La combinación de metas e información estructurada (que constituye el nicho cognitivo) en que se desenvuelven (las personas) (...) de una unidad organizativa, produce poderosas lealtades y visiones comunes de la situación que pueden llamarse *identificaciones organizacionales*. (Simon 1993: 159-160; paréntesis añadidos; cursivas en el original).

Pasemos a analizar brevemente la naturaleza de las estructuras neurobiológicas portadoras de cultura.

5.6.- El sustrato neurológico de los memes

Según Bergstrom (2002: 84) los patrones evolutivos de genes y memes pueden explicarse con un mismo modelo darwiniano; sin embargo, la diferencia estriba en que los ritmos de mutación y selección de los rasgos transmitidos son más rápidos en la evolución cultural. Esta acusada diferencia en los ritmos de mutación y selección, que hace de la replicación cultural un fenómeno extraordinariamente rápido en comparación con la biológica, se debe a la particular diversidad de procesos de transmisión y selección de los replicadores culturales³³⁷. La herencia genética pasa de progenitores a descendientes a través de la duplicación del ADN en la reproducción sexual, de la cual surge una nueva generación que integra genes de ambos progenitores en proporción idéntica en los animales diploides. Se trata de una propagación vertical, unidireccional y única, pues la generación de un cigoto es incompatible, con el mismo material genético, con la generación simultánea de otros cigotos (salvo el caso de división del cigoto en dos, que daría como resultado la generación de gemelos, y salvo sofisticadas técnicas de laboratorio que aquí no vienen al caso). Los mecanismos de evolución genética consisten en la copia de genotipos de una generación a otra, y las fuerzas de la evolución son las conocidas de *mutación* (que ocurre excepcionalmente con éxito), *deriva genética* y *selección natural*.

Sin embargo, en la evolución cultural los replicadores se transmiten instalándose en las redes neuronales encefálicas, con la particularidad de que el paso de unos a otros puede ser a través de transmisión vertical intergeneracional (con o sin parentesco biológico, uniparental o pluriparental, y en ambas direcciones), horizontal (p.ej. entre amigos, o entre maestro y aprendiz, todos de la misma generación), u oblicua (cuando el transmisor pertenece a una generación diferente del receptor). Además, la transmisión cultural puede realizarse de manera masiva y simultánea, de un transmisor a muchos receptores (transmisión magistral, como la de profesor a alumnos, o líder de opinión a sus seguidores), o a la inversa, esto es, masiva y simultáneamente de muchos transmisores a un sólo receptor (transmisión concertada, como el adoctrinamiento dentro de un grupo al nuevo afiliado). Los mecanismos de evolución cultural consisten en la copia de fenotipos de manera multidireccional e intergeneracional, y las fuerzas de

³³⁷ La explicación siguiente se basa en el texto de Cavalli-Sforza (2010[1996]: cap. 6). Un resumen aplicado puede encontrarse en Castro Nogueira, *et. al.* (2003: 164 y ss.)

la evolución son la *mutación cultural* (como los defectos en la transmisión, o la innovación), la *deriva cultural* (análoga a la deriva genética), la *variación dirigida* (que comprende tanto la educación planificada, la mejora secuencial de diseño, la adaptación consciente a nuevas circunstancias, la innovación, etc.), la *transmisión sesgada* (p.ej. la censura, la tergiversación o la discriminación informativa, la enseñanza sesgada, etc.), y la *selección natural* (a través de la influencia de los caracteres culturales en la eficacia biológica de los individuos, que condiciona su probabilidad de replicación biológica y cultural).

En el marco de estos procesos, el sujeto es la actividad mental consciente de un organismo humano enculturado, y ese ámbito de la existencia es cultural además de biológico. Todavía no se conoce bien cuáles son los sustratos materiales de la actividad mental, su unidad básica de actividad, o las relaciones entre actividades mentales asociativas y compositivas. En el s. XVIII el médico y filósofo Julien Offray Le Mettrie sostuvo que los pensamientos consisten en secreciones específicas del cerebro. Esta idea, hoy día descartada, formó parte de una teoría de la mente a la que todavía se daba crédito en tiempos de Darwin (él mismo se lo dio), de modo que la actividad mental era considerada un atributo de la materia como pueda ser la gravedad (véase Popper y Eccles 1984: 173). Actualmente tampoco conocemos bien la unidad orgánica de soporte biológico de la cultura, ni la unidad básica de replicación cultural que viaja en los cerebros y se transmite de unos a otros³³⁸.

Sí existe, no obstante, certeza científica de que la genética humana, en mucho mayor grado que todos los demás animales, produce organismos inacabados que necesariamente han de completarse mediante la asimilación de cultura. Esto se debe a las particulares capacidades mentales evolucionadas de la especie, las cuales hacen posible el aprendizaje humano, único en la naturaleza en su carácter tanto cuantitativo como cualitativo. La experiencia mental humana surge de una base de capacidades filogenéticas que, en interacción con el medio, producen el desarrollo y maduración de las capacidades ontogenéticas adaptadas para la supervivencia en ese medio (Edelman y

³³⁸ Dawkins (1993[1976]) denominó meme a una imprecisa e hipotética unidad básica de replicación cultural, y la define define como «una entidad capaz de ser transmitida de un cerebro a otro» (p. 256). Los ejemplos que cita son: «las modas en el vestir y en los regímenes alimentarios, las ceremonias y las costumbres, el arte y la arquitectura, la ingeniería y la tecnología» (*op. cit.* p. 248).

Tononi 1995). De este modo el cerebro -y el sujeto que este órgano es capaz de producir- no aparece embrionariamente programado, sino que desarrolla sus estructuras dentro de una programación genéticamente codificada, que necesita estimulación externa para completarse en un sentido favorable (o no) para la supervivencia del organismo³³⁹. Los implantes mentales establecidos durante la ontogénesis a través del aprendizaje formal o informal, son circuitos ontogenéticos adaptados al medio en que el organismo vive e interacciona. Los grupos de neuronas compiten en una suerte de darwinismo neural que fortalece los que aportan capacidades de supervivencia al organismo, mientras que aquellos otros irrelevantes o maladaptativos tienden a debilitarse o a causar problemas de disfuncionalidad³⁴⁰:

... muchos aspectos de la historia individual necesariamente interfieren en la conformación de cada una de las respuestas neurales. Y tal es la historia que selecciona, restringida por valores evolutivamente fijados, la dotación que ocasiona la impredecibilidad e individualidad que caracteriza lo humano. (Edelman y Tononi 1995: 99).

Para Edelman la frontera entre la física inanimada y la neurobiología viene trazada por la capacidad de la memoria. En la especie humana la identidad personal es el producto del acoplamiento de experiencias que se “imprimen” a nivel neurológico, generando un modelo conceptual del mundo, un nicho cognitivo. La sucesión de recuerdos y las transformaciones acaecidas entre ellos configuran otro modelo conceptual que recoge el paso del tiempo: la identidad personal. Así,

Cada individuo no está solamente sometido, como todos los demás sistemas materiales, a la segunda ley de la termodinámica, sino también a un conjunto con muchas capas de eventos, seleccionados e irreversibles, en su percepción y memoria. De hecho, los sistemas selectivos son irreversibles por naturaleza. (Edelman 1995: 203).

Los estudios de biofísica de la computación parten de la premisa de que las experiencias mentales tienen como soporte material un orden de procesos físicos a nivel molecular³⁴¹. En estos modelos computacionales de la mente la configuración física de

³³⁹ Afirma Edelman (1995: 201) en este sentido: «Por sujeto no entiendo solamente la individualidad que emerge de la genética o la inmunología, sino la individualidad personal que emerge de las interacciones sociales y en general con el medio».

³⁴⁰ Una síntesis sobre la historia de la tradición científica del darwinismo neural, o grupos neuronales de formación ontogenética, se encuentra en Rosenfield (1986).

³⁴¹ Un comentario del trabajo señero de Christof Koch, *Biophysics of computation* (1999), se encuentra en Scott (2000). Una entrevista académica en la que el científico explica su concepción de la organización de la consciencia en el cerebro humano está en Blackmore (2009: 125-136).

la materia, interconectada en un sistema biológico como es el cerebro, es el sustrato de las categorías, conceptos, creencias, normas y pautas de conducta, adscripciones, símbolos culturales, lenguaje, etc., de la psicología humana. Queda mucho aún por conocer, pero sí se admite, hoy por hoy, la hipótesis de que lo que llamamos aprendizaje tiene lugar en la neuroquímica y en la estructura de los circuitos neurales y de las complejas formas, e interconexiones, de las dendritas. Los mapas neuronales operativos y reforzados por la experiencia tienen configuraciones y estructuras estables, y constituyen sustratos neurológicos modales que integran percepción y cognición con valor psicológico y decisorio. Tales dispositivos modales se dan en el reino animal, y resultan funcionales sin necesidad de un módulo único especializado en la autoconciencia y en la transcripción del lenguaje de la física al de la mente. Aunque aún no sabemos explicar cómo se transforman los estímulos aferentes en señales electroquímicas, y fundamentalmente éstas en significado, sí existen indicios teóricos, ya sometidos a examen y experimentación, de que los contenidos mentales están respaldados por una tecnología de representación analógica capaz de producir *qualia* y significado sin necesidad de ser transcritos y expresados en un lenguaje unimodal específico³⁴². Estos *qualia* vienen pues a ser pensamientos producidos y registrados en sustrato orgánico (*embodied thoughts*) (Hardy-Vallée y Poirier 2006), y como sustratos corpóreos y materiales que son, pueden ser tratados igual que objetos físicos (y mentales) replicables, o sea, replicadores:

Existen pocas dudas de que el cerebro utiliza configuraciones activas de neuronas para representar las propiedades de las entidades y acontecimientos percibidas (...) un símbolo perceptivo es un registro de la activación neurológica que surge durante la percepción (...) Un símbolo perceptivo no es el registro del estado neurológico completo que subyace a la percepción. Más bien se trata solamente de una pequeña parte de ese estado, que representa un aspecto coherente del mismo (...) En vez de contener una completa representación

³⁴² Un enfoque doctrinal dentro de la ciencia cognitiva establece que el proceso de funcionamiento representacional, cognitivo y semántico, del cerebro en la percepción es analógico y que, sobre esta base filogenética (la única en primates no humanos y niños pequeños humanos que aún no han aprendido a hablar) ha evolucionado y se desarrolla ontogenéticamente un sistema de procesamiento no perceptual digital de la información. Éste último contiene una complejidad neurológica superior, y posibilita la facultad humana de pensamiento abstracto, la imaginación, la fantasía, etc. Un producto de la representación digital de la experiencia consciente es el lenguaje. Esto significa que existe una modalidad de pensamiento puramente perceptual, prelingüístico y precultural, y otro, propio únicamente de la especie humana, que integra parte de la información perceptual multimodal en un dispositivo de transcripción digital (que funciona con esquemas representacionales) cuyo producto es el lenguaje. En cualquier caso, es posible plantear la hipótesis de un sustrato neurológico estable tanto de los “pensamientos sin palabras”, como de las transcripciones lingüísticas de los mismos. Véanse al respecto Barsalou, *et. al.* (2003), y Barsalou (1999).

holística del estado neurológico perceptivo, un símbolo perceptivo contiene solo algún aspecto esquemático relevante. (Barsalou 1999: 582-583).

A diferencia de la replicación genética, en la que la unidad replicadora es una unidad física claramente preconfigurada, la replicación cultural puede ser explicada como procesos de adoctrinamiento, a través de variados procedimientos de aprendizaje social, que en sí mismos configuran las unidades replicadoras dentro del organismo. Éstas son generadas y almacenadas en los cerebros individuales, que actúan a modo de vehículos y entidades de propagación cultural (Boyd y Richerson 1987: 67). El soporte físico-neurológico de la actividad mental consciente, como es el caso de la percepción de objetos, intención de actuar, identificación de partes del cuerpo, sencillas operaciones de cálculo, u organización de la actividad mental en categorías de clasificación, ha empezado a ser estudiado con éxito en experimentos consistentes en mapear la actividad neurológica asociada a estas operaciones mentales. Aunque los métodos de estudio, modelos utilizados y la tecnología existentes en la actualidad aún adolecen de importantes limitaciones, sin embargo ya es posible realizar decodificaciones experimentales de los patrones de actividad neural relacionados con dichas operaciones (Smith 2013; Haynes, *et. al.* 2007; Urgesi, *et. al.* 2007). Por ejemplo, se ha escaneado con un éxito notable el espacio semántico de representaciones analógicas de miles de objetos y categorías de acción procesadas (de forma estructuralmente semejante) en distintos cerebros humanos (Huth, *et. al.* 2012). Otro experimento revela que los modelos de representación de la actividad neurológica de las áreas corticales visuales están empezando a ofrecer decodificaciones bastante precisas de los contenidos mentales surgidos de la estimulación inducida durante el sueño (Horikawa, *et. al.* 2013). El conocimiento del sustrato neurológico que soporta la actividad mental es ya un hecho, y cabe suponer que se producirán mayores avances en un futuro cercano. Pasemos a presentar una teoría del sujeto basada en los replicadores culturales, la identidad y el sentido humano de la propiedad.

5.7.- El Yo, la identidad, y el sentido humano de lo propio

La especie humana manifiesta una inherente propensión a reconocer el derecho de propiedad (Rossano, *et. al.* 2011), un fenómeno psicológica y conductualmente más complejo que la territorialidad característica de muchas otras especies. El

reconocimiento del derecho de propiedad, que hace a la humana una especie de propietarios, implica no sólo la capacidad de adquirir y defender lo adquirido sobre la base de alguna modalidad de función asimétrica de valor (Kahneman, *et. al.* 1991), sino además reconocerlo y respetarlo en otros individuos. Esta capacidad aporta una dimensión diferencial a la moralidad humana, y además resulta explicativa del apego e identificación con los replicadores constitutivos del patrimonio cultural común del grupo:

El hecho de tener conciencia de sí mismo significa también que los individuos y las organizaciones humanas existen rodeadas de un complejo ambiente de creencias, ideas, prejuicios, deseos y aversiones. Este medio ambiente mental domina a los hombres y sus organizaciones sociales de la misma manera que el medio ambiente físico domina el reino animal. (Boulding 1975[1953]: 111).

Singer y Kraft (2005) explican que las principales funciones de supervivencia que desempeña el cerebro humano, desde una perspectiva social, consisten en posibilitar la interacción con otros individuos, desplegar procesos de aprendizaje y manejar eficazmente las relaciones de cooperación y conflicto. Un cerebro de estas características aporta al organismo que lo contiene una inteligencia al servicio de la supervivencia y replicación integrado en grupos, o sea, una inteligencia social. Este argumento supone implícitamente una negación de la visión meramente objetual de los otros individuos, y acepta los principios psicológicos de agencia e intersubjetividad. Para los individuos humanos los demás individuos de su especie son considerados yoes dotados de vida mental en primera persona cuya conducta es intencional, de modo que la interacción con ellos presupone

... la capacidad de hacerse cargo de la situación mental y psíquica de los demás, reconocer sus deseos, intenciones y pensamientos y tenerlos en cuenta en las propias acciones. Podemos comprender cuanto nuestro prójimo hace y deja de hacer por la sencilla razón de que nuestro cerebro está en condiciones de construir una representación de la vida interior ajena, sin que ello tenga nada que ver con nuestro propio estado mental. (Singer y Kraft: 60).

La autoconciencia, el problema fuerte de las ciencias de la vida, tiene un protagonismo explicativo, junto con la intersubjetividad, en la evolución del modelo humano de supervivencia a través de una sociabilidad mental y representacional complejas (Turner 2000: 138-142). Así, el espejismo del yo cumple la función biológica

de permitir extrapolar el modelo subjetivo de agencia individual a otros individuos, y sobre esa base poder llevar a cabo interacciones sociales beneficiosas en ambientes que pueden ser arriesgados e inciertos³⁴³. De este modo, es posible integrar al organismo en una interacción grupal compleja que le aporte ventajas de supervivencia. Asimismo, la identificación con las emociones y estados mentales de los otros yoes posibilita la cooperación con desconocidos e incluso el altruismo incondicional. Como explica Sarah Blaffer (2009: 4 y 11), la diferencia social entre los humanos y sus parientes genéticos más próximos se debe a

... su habilidad para identificarse con sus semejantes y experimentar vicariamente sus emociones, especialmente las de sufrimiento (...) Esta capacidad para el entendimiento mutuo es la base evolutiva de nuestro comportamiento cooperativo.

En otro sentido, el yo empírico está “vestido” de una identidad individual integrada en grupos sociales (Jetten, *et. al.* 2010: 409). Esta identidad se compone de un autoconcepto histórico-narrativo dotado de atributos propios y propiedades, que se refieren tanto a lo propiamente somático, como a objetos externos (materiales e inmateriales) que el yo considera propios, con los que se identifica, habitúa y vincula (objetos y sujetos), y que vienen a ser atributos endo y exosomáticos, materiales y mentales, que constituyen una identidad psicológica que permanece en el tiempo (Zajonc 1968). Así, en una preclara intuición decimonónica William James sostuvo que el yo se autoconstruye agrupando y asociando atributos y “propiedades” con los que se identifica, y que se almacenan en la memoria emocional de lo “propio”. El modelo de “sí mismo” es un modelo de agregación selectiva por identificación, dentro de un mundo compuesto de sujetos y objetos:

Está claro que entre lo que un hombre llama *mí* y lo que simplemente llama *mío* es difícil de trazar una línea divisoria (...) En el sentido más amplio posible (...) el sí mismo de un hombre es la suma total de todo lo que *puede* llamar suyo, no sólo su cuerpo y sus poderes psíquicos, sino sus ropas y su casa, su esposa e hijos, sus ascendientes y amigos, su reputación y su trabajo, sus tierras y caballos, su yate y su cuenta bancaria. (James 1890: vol. 1. p. 291; cursivas en el original).

Según Metzinger (2003) la conciencia del sí mismo presupone una actividad neural de representación que coexiste con representaciones de otros objetos separados del yo. El

³⁴³ La base epigenética de estas capacidades relacionales reside en las habilidades de simular, representar e inferir las experiencias mentales de otros individuos (Singer y Lamm 2009).

yo es por tanto una representación unificada de atributos, recuerdos y objetos, tales como el cuerpo-mente y lo que podemos llamar ingredientes de un “yo periférico”. De este modo, el cerebro humano desarrolla epigenéticamente un modelo de representación del mundo y sus objetos, de sí mismo, y de otros yoes, y entre todos estos elementos el yo establece asociaciones y vinculaciones emocionales e instrumentales, que incluyen las relaciones objetuales de uso, de aversión y conflicto, así como de atracción y cooperación, o incumbencia. Las representaciones que lleva a cabo el cerebro con un significado relacional e identitario específico para el sujeto son explicadas del siguiente modo por la ciencia cognitiva:

Las unidades (de representación) con significado relevante incluyen objetos del mundo que son importantes para el individuo (p.ej. miembros de la familia, amigos, posesiones personales) y categorías (p.ej. tipologías naturales, artefactos culturales, acontecimientos), donde una *categoría* es un conjunto de individuos en el mundo o en la introspección. Una vez que un simulador llega a establecerse en la memoria para una categoría, ello ayuda a identificar a los objetos que se incluyen en esa categoría en ocasiones sucesivas, y posibilita realizar inferencias categoriales sobre ellos. (Barsalou 1999: 587; primer paréntesis añadido; cursivas en el original).

Una conclusión de esto es que la categoría de la propiedad consiste en un proceso neuropsicológico más, que opera a nivel inconsciente, y que capacita al individuo para una moralidad innata en torno al derecho de propiedad. Todo el entramado de relaciones entre el yo, otros yoes y los objetos que no son yoes tiene valor adaptativo, y podemos denominarlo *relaciones de propiedad*³⁴⁴.

Los individuos humanos se relacionan entre ellos y con los objetos externos de los que se valen para sobrevivir dentro de órdenes sociales que reconocen derechos de

³⁴⁴ Como explican Ramachandran y Blakeslee (1999: 312), para que el organismo no base su comportamiento exclusivamente en patrones de acción fijos y, de este modo, el yo pueda “tomar decisiones”, la representación de los objetos del mundo y de los otros yoes ha de estar dotada de emociones y programas cognitivos que permitan establecer una regulación estable y coherente de la vida social. Por ejemplo, los humanos reconocen relaciones de jerarquía, que a veces activan emociones de aversión y agresividad. Tales estados emotivos generalmente son inhibidos porque el modelo del yo está integrado en un modelo más amplio al que pertenece, y del que depende psicológica y somáticamente. Ese modelo de representación de los objetos y sujetos que componen el mundo percibido soporta relaciones de familiaridad y propiedad basadas en el reconocimiento de señales específicas de identificación de lo propio. Estas relaciones de propiedad vienen acompañadas de emociones y sentimientos de apego, afiliación y pertenencia:

... todas nuestras experiencias se pueden clasificar en categorías generales o símbolos (p.ej. personas o coches), así como ejemplares concretos o tipos (p.ej. Joe o *mi* coche). Nuestros experimentos parecen indicar que esta distinción (...) está profundamente arraigada en la arquitectura del cerebro. (Ramachandran y Blakeslee 1999: 218).

uso y dominio privilegiados, relaciones de intercambio, derechos de acceso individuales o comunitarios, y limitaciones al uso y disfrute de esos derechos (Fiske 1991, y 2002). Todo ello tiene un arraigo en los innatismos. Las bases de este complejo sistema de relaciones de propiedad son universales e instintivas, y sobre ellas se han erigido las culturas y las instituciones políticas y jurídico-económicas como fenómenos sociales epigenéticos³⁴⁵. La propia sociedad, como apostilla Ridley (1996: 6), «no fue inventada por el razonamiento humano (...) Es tan producto de nuestros genes como lo son nuestros cuerpos». En este mismo sentido Turner (2000: 1) puntualiza que:

La aparente distancia entre el moderno mundo sociocultural y la biología humana es una ilusión, porque la interacción y la organización social humanas son posibles sólo en virtud de la neuroanatomía y psicología homínida; y más fundamentalmente, la interacción y la organización están constreñidas por su legado biológico.

De este modo, el entramado de relaciones de propiedad al que aludimos está causado epigenéticamente por innatismos, que capacitan para la proyección externa del yo y la comprensión de otras mentes. Esos mismos innatismos explican las relaciones de apego e incumbencia (o de compromiso personal) hacia objetos y sujetos. El proceso neuropsicológico de la identificación desempeña un papel fundamental aquí:

... cuando jugamos al ajedrez, imaginamos que “somos” la reina y planeamos “su” próximo movimiento. Cuando hacemos esto, casi podemos sentir por un instante que hemos ocupado el cuerpo de la reina. Naturalmente, se podría argumentar que se trata sólo de una manera de hablar, que no estamos necesariamente asimilando la pieza de ajedrez en nuestra imagen corporal. Pero, ¿estamos seguros de que la lealtad de nuestra mente a nuestro cuerpo no es también una “manera de hablar”? ¿Qué ocurriría con nuestra respuesta dermogalvánica si alguien de repente golpeara a la reina? ¿Subiría de golpe como si nos hubieran pegado en el propio cuerpo? ¿Es posible que la tendencia normal a identificarnos con “nuestro” cuerpo y no con la pieza de ajedrez sea también cuestión de convenciones, aunque de carácter más duradero? ¿Podría un mecanismo semejante ser responsable también de la simpatía y el cariño que sentimos por un amigo íntimo, un cónyuge o un hijo que es, literalmente, un producto de nuestro propio cuerpo? (Ramachandran y Blakeslee 1999: 312-313).

³⁴⁵ Margaret Gruter (1991) señala que los sistemas jurídicos y las regulaciones de la propiedad resultan de la formalización cultural de tradiciones y normas sociales informales, que a su vez son el resultado de la coevolución de emociones sociales y las propias normas y tradiciones. En este sentido, Jones y Goldsmith (2005) han investigado sobre los fundamentos biológicos de la justicia. Como subraya Marcel Mauss (1971[1950]: 18) refiriéndose al sentido profundo de la justicia en el ser humano: «Los hombres han sabido comprometer su honor y su nombre mucho antes de saber firmar.»

Tratamos seguidamente la subjetivación como un proceso de identificación consistente en la búsqueda y asentamiento histórico de una identidad dentro de un ambiente social y cultural.

5.7.1.- ¿Quién soy yo? El yo-propietario a la búsqueda de su identidad

El antropólogo Alfred I. Hallowell sostuvo que el gran desafío de la sociobiología es explicar la evolución de la conciencia reflexiva y la identidad personal en el ser humano³⁴⁶. Según Hallowell sólo las especies culturales o proto culturales tienen éxito en llevar a cabo rápidas adaptaciones del comportamiento a las condiciones locales sin necesidad de alteraciones en el genoma. Se trata de una habilidad que podríamos denominar *camaleonismo fenotípico*, y en la especie humana se fundamenta principalmente en su desarrollada neuroplasticidad adaptativa. La neuroplasticidad humana está causada por su genética, pero la causa próxima que permite la flexibilidad adaptativa a los entornos locales es, según Hallowell, la psicología del yo y su relación con el sistema normativo-cultural en el que se integra el individuo. Los animales sociales y culturales se adaptan a un orden social de prescripciones normativas porque su sistema nervioso experimenta una tendencia instintiva a la aceptación conformista de la influencia social. Para que las normas sean interiorizadas es preciso que el organismo disponga de un modelo que supervise la adecuación del comportamiento del individuo a las prescripciones normativas de su entorno social. Ese modelo es la conciencia reflexiva, o yo, y cuando el sistema nervioso realiza la función de autoevaluación (y evaluación de los otros) de manera automática y funcional, se puede decir que el individuo ha internalizado las normas. No obstante, la internalización normativa, que es una capacidad filogenética de la especie, consiste en la valoración intrínseca de determinados comportamientos (o creencias, actitudes, valores, normas, adscripciones y vinculaciones; un conjunto de información que podría sintetizarse como creencias y preferencias), a los que se asocia una valencia emocional positiva, y ello a pesar de que, en ciertas ocasiones, puedan ser la causa una reducción de la eficacia y/o del bienestar del individuo (Bowles y Gintis 2003: 439).

³⁴⁶ Lo aquí tratado sobre Hallowell procede del estudio monográfico de Barkow (1978).

La internalización del ambiente social, o “mapa cognitivo”, requiere un cerebro complejo, capacitado no sólo para el aprendizaje social, sino igualmente para la lectura de otras mentes y para la adscripción de valores a reglas, símbolos, actitudes y comportamientos. Todo ello aporta estabilidad relacional y cultural al sujeto, que de otro modo, y dispuesto de un cerebro capacitado para adaptarse a una gran diversidad de entornos sociales y culturales (y a sus respectivas identidades), podría caer en faltas de coherencia interna, y resultar impredecible para sí mismo y para los demás³⁴⁷. Obviamente, esto supondría un problema adaptativo en un animal que obtiene eficacia de la interacción social. La psicología humana ha evolucionado en el sentido de evitar la disgregación del autoconcepto, y opera activamente para mantener unidad y coherencia subjetivas.

Desde esta perspectiva, los estudios realizados sobre la identidad personal enfatizan su carácter reductor de indeterminación, y por lo tanto del estrés derivado de la inestabilidad identitaria. El asunto guarda relación con el problema filosófico de los múltiples yo es señalado entre otros por Parfit (2004[1984]) y en Elster (1997)³⁴⁸. Un organismo social dotado de *camaleonismo fenotípico* y capacidades cognitivas avanzadas puede sentir angustia existencial causada por la indeterminación de su autoconcepto y la consecuente incertidumbre, tanto de su comportamiento (público y privado) como de su adscripción a grupos sociales. Asimismo, la identidad se asocia a atributos físicos, valores y comportamientos socialmente prescritos, de modo que una vez interiorizada la identidad existen procesos psicológicos innatos que autorregulan la fidelidad y la defensa pro-activa (incluso frente a terceros) de tales rasgos. Nuevamente el mecanismo psicológico autorregulador es aversivo, y consiste en el estrés psíquico que produce la percepción de una amenaza a la identidad, interna o externa, dado que los individuos se esfuerzan por preservar su identidad (Thomas 1996). De este modo, en el fenómeno de la identificación se producen externalidades (positivas y negativas), que son internalizadas a través de la aversión automática de las desviaciones respecto a la

³⁴⁷ Estos comportamientos son considerados patológicos, y pueden estar causados por enfermedades mentales graves como las psicosis, consideradas incapacitantes en diverso grado en función del ambiente psicosocial predominante. Véase Molina (2012) para una investigación reciente al respecto.

³⁴⁸ Una de las derivaciones del problema de los múltiples yo es el cambio errático de preferencias y la debilidad de la voluntad (Elster 2000[1979], y 1997). Para evitarlo, la psicología humana, como la propia evolución, tiene un sesgo conservador, tal y como señala Schelling (1985: 358 y 362) en su descripción de las inercias psicológicas, y la aversión a la incertidumbre de los cambios a futuro en el propio auto-concepto.

categoría socialmente construida³⁴⁹. Téngase en cuenta que desde una perspectiva de neuropsicología clínica la identidad tiene una función adaptativa, dado que es un factor homeostático del organismo humano, y en la medida que reduce el estrés asociado a la incertidumbre conductual y relacional, favorece la adaptación al medio social y la eficacia biológica (Jetten, *et. al.* 2010: 409). Igualmente ocurre con el fenómeno de la disonancia cognitiva, un proceso automático del cerebro programado para evitar la incoherencia entre el autoconcepto, los pensamientos, sentimientos y comportamientos empíricos³⁵⁰.

La teoría de Hallowell conlleva en este sentido dos implicaciones. Por un lado, el yo es un producto evolutivo funcional, que al posibilitar el cotejo permanente de la representación del individuo en otra representación más amplia y compleja (que incluye a otros individuos y un sistema social-cultural de reglas de comportamiento), confiere una estrategia de supervivencia evolucionada. El individuo que es además un yo dispone de un sistema autorregulado y autoevaluado de existencia. Esta estrategia aporta eficacia a los individuos capacitados (genéticamente) para interiorizar normas de un ambiente cultural. Tal conclusión, propia de un enfoque adaptacionista del cerebro humano, no fue fundamentada evolutivamente por Hallowell, si bien sí se ha hecho con posterioridad y cuenta con acreditada solvencia teórica³⁵¹. En segundo lugar, el yo consiste en una capacidad psicológica evolucionada para “absorber” cultura durante la ontogénesis, y eso implica la operatividad de un proceso psicológico de identificación del organismo con estereotipos culturales³⁵². El proto-yo en sí mismo es un proceso de

³⁴⁹ Akerlof y Kranton (2000: 717) mencionan el ejemplo de la identidad de género: masculino y femenino. Así, cuando se está adscrito a la identidad masculina el vestirse con ropa culturalmente asociada a la identidad femenina produce malestar psíquico en el propio sujeto (caso de la emoción de la vergüenza), e incluso en otras personas (activando emociones como la ira o el odio). Estas externalidades (del sujeto y/o de otros sujetos) son la explicación de las crisis (psicopatológicas) de identidad, así como de los procesos de exclusión, segregación y estigmatización social.

³⁵⁰ La investigación de referencia en este tema es Festinger (1957). Un análisis psicológico actual puede encontrarse en Eakin (2001) y en Norton, *et. al.* (2004).

³⁵¹ En el enfoque adaptacionista el cerebro se presenta como una herramienta del comportamiento altamente versátil, lo cual no equivale a la hipótesis empirista tradicional de la *tabula rasa* (véase Sober 1996: 341).

³⁵² Don Ross (2008: 133) define la identidad como una «disposición humana universal a construir *yoes* (*selves*), y a reforzar tales construcciones en la interacción social» (cursivas en el original). Akerlof y Kranton (2000: 718) la definen como «auto-imagen de la persona así como sus categorías asociadas», consistiendo las categorías en modelos ideales estereotipados de comportamiento, apariencia física, pensamiento, valores y otros atributos socialmente prescritos. Queda claro pues que la identidad, enfocada hacia el auto-concepto individual o *ego*, es una construcción social que evoluciona por replicación.

recepción y representación de sensaciones y sentimientos relacionados con la homeostasis corporal, sin contenido cultural alguno (Damasio 2001). Ese proceso neurológico necesita “llenarse” de información, símbolos y de reglas (objetivos abstractos y específicos, sistema de creencias y compendios de técnicas y conocimientos tecnológicos). Una vez realizado este “llenado por identificación”, el organismo ya dispone de una identidad personal y social que incluye la evaluación de la información, semántica y reglas del entorno social en el que se desenvuelve. La tipología de conductas interiorizadas en el proceso de aprendizaje por identificación llega a programarse neurológicamente con tal fortaleza durante la ontogénesis, que se asemeja a los instintos “implantados”. Así ocurre, por ejemplo, con las reglas y valores interiorizados en la socialización a través de la herencia cultural, y su capacidad de generar emociones que permiten segmentar las poblaciones en grupos delimitados: nosotros *versus* ellos (Thornhill, *et. al.* 1997: 243 y ss.)

En efecto, el yo es un modelo cognitivo de interpretación de la información social-relacional, y esa interpretación se lleva a cabo con habilidades lógicas y emociones. La identidad personal puede entenderse en la línea de lo que Sen (1977, y 1985) denomina “compromiso” con una serie de reglas y valores, que operan como principios sobre los que se establecen preferencias “acotadas” (por marcos preferenciales de orden superior). Esto no obsta para que tales “metapreferencias”, y el compromiso pre-racional en el que se sustentan, puedan jugar en contra de los intereses del individuo a corto plazo (Davis 2006: 378). Asimismo es compatible con un proceso psicológico especializado en grabar creencias con valencia emotiva que categorizan automáticamente la interacción con identidades rivales como juegos de suma cero (Norton y Sommers 2011). Esto consiste en lo que Akerlof y Kranton (2000: 737 y ss.) asocian a la idea de “trampa de la identidad”, en la que la fidelidad a los rasgos identitarios asociados a la exclusión social y a la pobreza, como ocurre con las culturas de ciertas minorías étnicas en los países desarrollados, es una estrategia de confrontación identitaria con un valor intrínseco altamente costoso para el individuo y su grupo étnico-cultural³⁵³. Lógicamente, la identidad personal no es una elección

³⁵³ Los autores explican que en este fenómeno la identidad se refuerza por oposición a la identidad social dominante. De modo que los comportamientos y valores propios de la identidad social de las minorías vienen definidos no en aspectos favorables al interés del grupo, sino que se consigue la cohesión del grupo asumiendo comportamientos y valores que representan *de facto* la negación de aquéllos que son

utilitaria, sino en gran medida un conjunto de rasgos fenotípicos interiorizados por transmisión cultural, que se graban en la memoria emocional del individuo, y que producen acciones y reacciones no deliberadas³⁵⁴. Una vez implantada la identidad, ésta pasa a ser un *modelo cognitivo* estable con el que el sujeto asimila y da significado tanto a sí mismo como a la información que recibe de la interacción social (Pinker 2004: 246).

Una conclusión importante derivada del análisis previo es que el yo dotado de significado autorreferencial no sólo es una perspectiva desde la cual se observa pasivamente el acontecer en un proceso de “adherencia” (o aprehensión) aleatoria de información (como en la teoría lockeana), sino un filtro discriminador que enjuicia y valora; o sea, el yo identificado es como un cedazo metafórico que criba la información aferente con emociones y juicios sobre sí mismo y sobre lo que observa. El ego necesita *ser* “algo” para existir, creer que es un agente que lleva el control de sí mismo, ser un sujeto con significado social, y para ello construye (asimila coherentemente) una identidad personal a través de la cual se construye un sujeto. Lo hace identificándose con conceptos y valores (replicadores) que hace “suyos” a través de una obsesiva identificación emocional, laboriosa y constructiva, tal y como expresa -impremeditadamente- Isaiah Berlin cuando aclama que el sentido positivo del concepto libertad se deriva del impetuoso deseo que tienen los individuos de tener el control de sí mismos, lo cual incluye la necesidad de «ser alguien y no nadie» (Berlin 1978: 140).

En su dimensión social, el yo no sólo es algo para sí mismo, sino que requiere ser algo para los demás, y serlo dentro de categorías de pertenencia en torno a las que se estructuran grupos de otros yoes. Así, Tajfel (1972: 292) ofrece una definición de la identidad social engarzada en una semántica emocional de afiliación y lealtad:

definitorios de la mayoría dominante. Esto puede tener efectos nocivos en la replicación biológica y cultural de esa identidad, y dado que las identidades sociales pueden asemejarse a la idea de “nichos ecológicos”, el refuerzo identitario puede conducir al debilitamiento y extinción de los grupos culturales si se aferran a comportamientos y valores perjudiciales para sus individuos, como pueden ser, por ejemplo, la violencia o el consumo de estupefacientes.

³⁵⁴ Aunque Akerlof y Kranton (2000: 726-727) admiten cierta capacidad de elección por parte del individuo respecto a su identidad, sin embargo reconocen que esa capacidad es limitada, dado que existen influencias de otros individuos, así como del ambiente cultural y del adoctrinamiento, que escapan de su control, y ante los cuales el yo es una disposición receptiva. Véase al respecto igualmente Richerson y Boyd (1993).

... el conocimiento del individuo de que pertenece a ciertos grupos sociales, unido a una significación emotiva y de valor para sí mismo derivada de su pertenencia al grupo.

En esta tarea identificatoria el cerebro establece asociaciones estables con experiencias propias y ajenas, así como con objetos sin yo (que son objetos, símbolos, posesiones, y los propios pensamientos y acciones, etc.), con todo lo cual se vincula en una red de relaciones idiosincrásicas de incumbencia y propiedad que se cree están bajo “su” control (Klein, *et. al.* 2002: 127). Hull y Zacher (2007: 75) afirman en este aspecto que

... el sentido de lo que somos nosotros mismos surge en relación con la identidad de otros; algunas veces en armonía con ellos, otras veces enfrentado a ellos, pero siempre en relación a ellos.

El yo identificado con un estereotipo cultural ya no es un organismo conductualmente adaptable a cualquier ecosistema social o ideológico, sino que la enculturación favorece la estabilidad propia de lo que Edgar Morin denomina *sujetamiento*, o sujeto sujetado; una capacidad humana derivada de los instintos sociales que caracterizan su naturaleza psicológica propensa a la influencia social:

Ser sujetado no significa ser sojuzgado desde el exterior, como un prisionero o un esclavo. Significa que una potencia subjetiva más fuerte se impone en el centro del dispositivo lógico egocéntrico y, literalmente, subyuga al individuo, que entonces se encuentra poseído en el interior de sí mismo. De este modo, el sujeto (en el sentido autónomo del término) puede devenir sujeto (en el sentido dependiente del término) cuando el Superyó del Estado, de la Patria, del Dios o del Jefe se impone en el interior del dispositivo lógico de inclusión (...) Podemos ser poseídos subjetivamente por un Dios, un Mito, una Idea; y son esta idea y este mito los que, inscritos como un virus en el dispositivo lógico egocéntrico, nos darán órdenes imperiosamente creyendo servirlos voluntariamente. (Morin 2004: 84).

Tras todo ello, la identidad acaba convirtiéndose en un *a priori* cuando en el sistema límbico se graban y refuerzan emociones vinculadas a comportamientos y atributos simbólico-culturales. Cuando el individuo se siente emocionalmente identificado con un determinado imaginario cultural, o conjunto coherente de replicadores culturales, su comportamiento se aparta del sistema de incentivos consecuencialista propio de la elección racional, y se convierte en un fin intrínseco, un “modo de ser”. Así, como indican Hoffman, *et. al.* (1998: 350), cada comportamiento contribuye al refuerzo de los anteriores, incluso los realizados en microexperimentos sociales de laboratorio. Como los comportamientos que no son experiencias aisladas

forman parte de una secuencia vital, así las respuestas a los estímulos vienen a ser estereotipadas y coherentes con un conjunto de normas de conducta. Hume y Mill lo explican con un lenguaje filosófico ilustrado propio de su época, pero suficientemente esclarecedor:

Cuando un hombre está poseído de una alta autoestima de su lugar y condición en el mundo, naturalmente se verá obligado a actuar en consecuencia y evitará realizar actos de bajeza que puedan hundirlo por debajo de la imagen de sí mismo que él se ha formado en su propia fantasía. (Hume 1995[1741-42]: 52).

Ningún ser humano inteligente admitiría convertirse en un necio, ninguna persona culta querría ser un ignorante, ninguna persona con sentimientos y conciencia querría ser egoísta y depravada, aun cuando se le persuadiera de que el necio, el ignorante o el sinvergüenza pudieran estar más satisfechos con su suerte que ellos con la suya. (Mill 1991[1863]: 48).

Akerlof y Kranton (2005: 9 y 17) inciden en estas adscripciones no-utilitarias de la identidad proponiendo un ejemplo contemporáneo, el de los marines del ejército de los EE.UU.:

El propósito de la Armada en West Point es cambiar las preferencias de los cadetes. Se pretende inculcar motivaciones no-económicas en los cadetes, de manera que lleguen a tener las mismas metas que las de la Armada de los EE.UU. (...) el objetivo de West Point es cambiar la *identidad* de los cadetes, de modo que lleguen a pensar acerca de sí mismos, sobre todo y más, como oficiales de la Armada de los EE.UU. (...) El oficial debe obedecer las reglas de la organización y cumple órdenes procedentes de la cadena de mando. No debe seguir esas órdenes pasivamente, sino que debería tener “fe en el sistema”. De hecho “perder la fe en el sistema supone colocar el ego antes que el espíritu de servicio”, y eso es una traición al lema de la Fuerza Aérea que dice “servicio antes que tú mismo”. (Cursivas y comillas en el original).

Tanto el hombre de alta condición social de Hume, como el virtuoso de Mill, y el soldado de Akerlof y Kranton, son sujetos sujetos a una identidad fijada culturalmente en sus dispositivos orgánicos de consciencia, evaluación y conducta. Los tres son organismos construidos por replicadores genéticos en interacción con un determinado medio natural y social, y que además portan replicadores culturales interiorizados a través de programaciones ontogenéticas para la supervivencia en sociedad. La programación ontogenética, a diferencia del genoma único e individual, es un producto cultural estandarizado del que disponen los individuos que integran una determinada cultura. Consecuentemente, el comportamiento altruista de la pertenencia cultural es el

resultado de una evolución por selección cultural y grupal, en los mencionados sentidos débil y fuerte.

5.8.- Conclusiones

Si un individuo recién nacido tiene mayor esperanza de vida y de reproducción dentro de un determinado grupo humano que en otro se puede sostener que, *cæteris paribus* la genética, han de estar operando otros factores de selección adicionales a los genes. El ecosistema cultural se muestra un elemento determinante en la supervivencia y reproducción biológica con independencia del genoma. Los grupos humanos se caracterizan por poseer culturas, y se puede sostener que las fuerzas selectivas del grupo residen básicamente en su cultura peculiar. Incluso admitiendo la influencia de los instintos parentales de selección inclusiva en el comportamiento, una misma familia ubicada en un ambiente cultural u otro tiene diferentes probabilidades de supervivencia y replicación, así como diferentes pautas de comportamiento endo y exofamiliar. Así, dos gemelos homocigóticos, que comparten el 100% de su dotación genética (o sea, son técnicamente clones), que sean separados nada más nacer y enculturados en distintos grupos u ambientes, manifiestan empíricamente diferentes esperanzas de vida y de éxito reproductivo asociadas a las diferencias que, en creencias y estilos de vida, determinan una diferente enculturación (Cavalli-Sforza 2010[1996]: cap. 6). Esto pone de manifiesto la presencia de dinámicas selectivas inherentes asociadas al grupo y a la cultura, es decir, un patrón de selección cultural-grupal que opera paralelamente a la selección genética.

Las paradojas evolutivas de la cultura, como por ejemplo la supervivencia y amplia difusión de creencias irracionales, o las demostraciones de compromiso con el grupo y sus creencias, frecuentemente llevadas a cabo a través de sacrificios personales altamente costosos, son en realidad paradojas de la identidad. La cultura aporta identidad a los individuos, y esa identidad compartida es la responsable de los procesos humanos de selección grupal. Participar en una cultura conlleva costes y beneficios de supervivencia, y afectan a los grupos como conjuntos poblacionales, y no sólo como individuos yuxtapuestos. La especie humana habita ecosistemas espaciales y nichos cognitivos en los que se adapta y sobrevive. En ellos se pone de manifiesto una

dimensión específica del comportamiento altruista, y permiten explicar teóricamente la evolución del altruismo.

En el presente capítulo se ha estudiado cómo los sistemas de replicación cultural, en interacción con un cerebro dotado de gran neuroplasticidad, posibilitan aumentar la convergencia fenotípica en los grupos humanos, y de ello la cooperación social a través de procesos de selección grupal. La diseminación de los replicadores culturales equivale a una definición de patrones de comportamiento constitutivos de puntos focales en casos de equilibrios múltiples, equilibrios de Nash no cooperativos, o procesos evolutivos del tipo carrera armamentista.

La prosocialidad humana está adaptada para aprovechar ventajas de supervivencia derivadas de la interacción espontánea en, así como la construcción de, ambientes relacionales generadores de sinergias y otras ventajas de cooperación. Tales ecosistemas culturales incrementan la eficacia del individuo, aun operando éste dentro de una vida regulada y limitada en cuanto a posibilidades de acción. Esta cesión a la estrategia cooperativa es una manifestación de comportamiento altruista que permite al individuo formar parte de un grupo fenotípicamente homogéneo, con un determinado grado de comunidad en los objetivos vitales, así como en el disfrute y aprovechamiento de bienes también comunes. La disposición procultural innata contenida en el instinto de conformismo se sustancia en la receptividad a la influencia social a través de procesos de adoctrinamiento activos y pasivos, que explican modelos de selección grupal en sentido fuerte y débil respectivamente. Una vez interiorizados los replicadores culturales en el adoctrinamiento, los grupos dotados de fenotipos homogéneos disponen de una mentalidad grupal más aproximada, que posibilita el éxito de los procesos sociales de elección de metas comunes, así como de la acción colectiva para la consecución de dichas metas de destino común. Puede hablarse de una mentalidad perteneciente al grupo en cuanto a grupo, reforzada por el comportamiento de sus miembros integrantes:

La mentalidad grupal descansa en un tipo de intencionalidad colectiva en la que todos los integrantes del grupo participan, tanto cumpliendo como reforzando las normas que definen al grupo, y estabilizan sus actividades cooperativas. En este aspecto, indudablemente la selección cultural grupal desempeñó un papel importante. (Tomasello, *et. al.* 2012: 685).

La desarrollada capacidad para la plasticidad fenotípica es un rasgo definitorio de la naturaleza humana que hemos denominado camaleonismo fenotípico. Este rasgo hace al hombre un animal esencialmente cultural, y por ello adaptado para sobrevivir culturalmente y esforzarse por transformar y propagar con éxito el conocimiento que porta en su sistema nervioso central. Como sostiene Bartra (2006: 42), la ontogénesis del cerebro no es un proceso determinista en su definición y estructura, puesto que la plasticidad cerebral consiste en los procesos de neurogénesis y desarrollo de circuitos que no están predeterminados genéticamente en un sentido estricto, y que de no producirse con la exposición al medio social no aparecen o no maduran. La naturaleza humana, de este modo, sólo se despliega epigenéticamente con éxito en sociedad, a través de la interacción social y la asimilación y propagación de cultura. En consecuencia, nunca hubo un estado de naturaleza “salvaje” sin sociedad humana, por muy elemental que ésta fuera, y nunca ha habido sociedad humana sin cultura. El individuo humano, por eso, no sólo está adaptado para saber *que es* (dispone de consciencia reflexiva), sino además para saber *qué es* (dispone de identidad personal y social), es decir, persigue inconscientemente ser un actor cultural identificado con rasgos culturales que le aproximan o alejan de otros individuos, por lo que la tercera adaptación derivada de la plasticidad fenotípica es saber *a qué se parece* (si dispone o no de afines y de afiliación a grupos en los que coopera).

Igualmente destacable es que los replicadores culturales constitutivos de la identidad personal conectan el binomio de supervivencia naturaleza – cultura, y tienen una derivada que contribuye a reforzar el valor adaptativo de los innatismos prosociales de la especie: el cerebro humano asimila cultura a través de la imitación e identificación en procesos de aprendizaje social. Esta técnica epigenética de interiorización de rasgos culturales y adaptación al medio social y cultural no es neutral en cuanto a los efectos en la motivación y el comportamiento del individuo en sus relaciones sociales. Hemos explicado cómo la plasticidad fenotípica de la especie humana construye identificaciones e identidades ulteriores durante la ontogénesis, que influyen decisivamente en la naturaleza de las relaciones (de cooperación, conflicto, o indiferencia pasiva) del individuo con los demás individuos de una población dada. El *pool* de replicadores biológicos y culturales que constituye el organismo delimita las

afinidades y semejanzas tras las que operan innatismos orientados a la eficacia inclusiva exofamiliar.

Consecuentemente, el altruismo es una conducta que, en la especie humana, no se acota de forma exclusiva al ámbito genético-parental, puesto que la naturaleza dual de la especie (biología-cultura) amplía el alcance de combinaciones de replicadores que permiten la activación de innatismos de causalidad próxima motivadores de actos altruistas. La evolución de las estrategias culturales consecuencia de la evolución de los replicadores culturales posee capacidad explicativa del grado de convergencia fenotípica de una población, y consecuentemente de su nivel de semejanza y de unidad de destino en los replicadores. Esta unidad de destino, en este caso de naturaleza cultural, resulta igualmente explicativa de los niveles de cooperación social. Ello se debe a que la semejanza fenotípica activa procesos psicológicos de identificación que motivan a comportamientos altruistas hacia los portadores de rasgos definitorios de esa identidad asumida como propia.

La identidad humana se asienta en el fenotipo, y es una identidad mixta genética y cultural señalizable a través de marcadores que delatan los replicadores que porta el individuo, y que asimismo regulan la expresión de comportamientos. Los marcadores de identidad, fundamentalmente culturales en la especie humana, facilitan la identificación cultural de los individuos, reduciendo de este modo la incertidumbre en la interacción típica de órdenes sociales complejos característicos de la especie³⁵⁵. Estos mismos marcadores sirven además como motivadores de la cohesión social y demostración de lealtad al grupo. Se ha defendido la hipótesis de que los marcadores de identidad cultural son operativos por ser consecuencia de exadaptación de los instintos inclusivos y de pertenencia previamente adaptados para la *kin selection*. Actúan por tanto como señuelos de los instintos asociados a los comportamientos de eficacia inclusiva. Los señuelos culturales que son los marcadores permiten que muchas adaptaciones prosociales del enclave familiar o dunbariano, se manifiesten operativamente en

³⁵⁵ Como establece Ross (2008: 133): «la construcción del yo es solamente el primer (y necesario) aspecto para el logro de la coordinación de grandes números.» Igualmente, Ramachandran y Blakeslee (1999: 317) concluyen que: «aquí tenemos la mayor ironía: que el yo, que casi por definición es algo completamente privado, es en gran medida un artefacto social, un cuento que inventamos para los demás.»

agregados relacionales más amplios y complejos.

Los implantes culturales son programaciones epigenéticas del cerebro que resultan posibles debido a los innatismos adaptados del conformismo. El individuo conformista asimila reglas y creencias culturales, y no optimiza sistemáticamente, sino que sigue normas grupales generadoras de equilibrios cooperativos dentro de la especie o del grupo. La herencia dual humana (genética y cultural) es una solución adaptativa que permite además optimizar recursos cognoscitivos. En este sentido, la herencia cultural nos sitúa ante la hipótesis de optimización ecológica dentro de grupos culturales, y no de unidades individuales. El individuo no es una pieza única de la especie humana, sino una réplica enculturada más o menos exacta dentro del grupo y de la especie. De ello que su operativa de supervivencia se encuadre en procesos de selección grupal compatibles con la selección en varios niveles. Aquí ofrecemos una teoría cultural del altruismo, explicado por restricciones culturales a la libertad de acción individual, unido a comportamientos progrupales. De este modo, aunque el altruismo puede tener una explicación puramente genética, sin embargo no necesita ser un rasgo codificado genéticamente para existir, dado que es suficiente con que se trate de un subproducto evolutivo que “viaja como autoestopista” de la previa adaptación prosocial y procultural (ésta sí, codificada en el genoma) del conformismo.

Actualmente existen indicios experimentales que permiten proponer una hipótesis sobre el sustrato material de los memes culturales como unidades de análisis, así como de la replicación a través de su implantación biofísica en los cerebros. Los replicadores de la evolución cultural se transmiten instalándose y dando forma a las redes neuronales encefálicas, mediante procesos de transmisión más diversos y rápidos que los de la replicación genética. Lo que llamamos aprendizaje cultural tiene lugar en la neuroquímica y el diseño de estructuras de circuitos neuronales, que se constituyen en el transcurso de la ontogénesis dentro del medio social y cultural. De confirmarse esta posibilidad, podría plantearse la herramienta teórica de un coeficiente de relación que incluyera los replicadores genéticos y culturales, o sea, una “triple hélice” que puede denominarse teoría “geme” (gene – meme).

Una consecuencia de lo anterior es que se abre la posibilidad de caracterizar el yo – sujeto como un entramado psicológico de atributos identitarios unificados por sentimientos de incumbencia, que también pueden denominarse sentimientos de “propiedad”. La subjetivación vendría a ser un punto focal de la mente en su afán epigenético de preservar o propagar los replicadores que constituyen el organismo. El yo, problema fuerte del materialismo, se explica como producto evolutivo funcional, como ya se hizo en el capítulo 2 de esta investigación. En el presente capítulo además hemos avanzado hacia el análisis del yo empírico, insertado socialmente a través de la implantación de replicadores culturales en un cerebro, y que se adhiere a ellos a través de lo que llamamos identificación psicológica, o más sencillamente “compromiso” de propietario. El compromiso con la identidad se expresa con la activación de sentimientos y procesos cognitivos cuya expresión comportamental es lo que comúnmente llamamos propiedad. De este modo, el sujeto puede ser explicado por una programación por identificación, o modelo cognitivo identificatorio con objetos físicos y mentales. Un epifenómeno así generado constituye una corriente de acción no aleatoria, sino regulada, que además tiene el refuerzo de una regulación social que puede ser de un carácter más activo o pasivo, pero que en cualquier caso engarza con que el yo infiere las actitudes y juicios de otros yoes sobre su comportamiento. Este es el origen de la moralidad humana y de las instituciones sociales epigenéticas de la moralidad, como son las costumbres, tradiciones y sistemas jurídicos.

El cerebro humano desarrolla epigenéticamente, y en interacción con el medio cultural, un modelo de representación del mundo y sus objetos, de sí mismo, y de otros yoes, y entre todos estos elementos se establecen asociaciones y vinculaciones emocionales e instrumentales, que incluyen, entre otras, las relaciones objetuales de uso, aversión y conflicto, así como de atracción y cooperación, o incumbencia. A diferencia de la territorialidad, la propiedad puede encajarse en este modelo de sujeto como un epifenómeno del yo, surgido de la interacción entre biología y cultura, que contribuye a reforzar la sociabilidad humana. El yo es, de este modo, una adaptación pro-cultural y pro-cooperativa, que ha permitido a la especie humana cooperar dentro órdenes sociales complejos, y en dinámicas de selección cultural – grupal.

Así pues, el modelo de yo-propietario es un modelo de agregación por identificación selectiva, que incluye un yo-periférico (exosomático) de atributos físicos y mentales. El sujeto así expresado en un modelo emotivo-cognitivo-operativo tiene un carácter integrador y adquisitivo, coherente con la realidad de un proyecto de replicación no-acabado, que pasa su vida adquiriendo y propagando cualidades, conocimientos e identidad. La subjetivación puede entenderse pues como un proceso histórico-narrativo de búsqueda e interiorización de una identidad personal dentro de un ambiente social y cultural. El yo programado por identificación es un modelo autorregulado de comportamiento, que sigue instrucciones de supervivencia de acuerdo a un plan coherente. Este yo enculturado no es capaz de renunciar a su identidad aun en situaciones cuya renuncia le aportaría ventaja neta de supervivencia. Y es que renunciar a ello es, en buena medida, renunciar a sí mismo y negar su propia existencia. Se trata pues de un sujeto que, como explican Thaler (2000) y Herrmann-Pillath (1994), no puede consistir en un programa de acción intrínsecamente utilitarista.

La teoría del sujeto como nexo identitario aquí planteada conecta con la explicación del comportamiento altruista, dado que la identificación compartida con un patrimonio de replicadores comunes explica un altruismo identitario en pos de la supervivencia de ese *pool* de replicadores, por encima de los intereses concretos y puntuales del organismo individual a corto plazo.

Hemos defendido que cuando la psique se identifica con un determinado estereotipo cultural, el individuo tiende a cooperar en mayor medida con los afines con quienes resulta más precisa la identificación. La causación próxima de esta operativa identificatoria es la empatía, así como otros procesos mentales y de acción común relacionados con la intersubjetividad. En el siguiente capítulo se analizan la cualidades filo y ontogenéticas que hacen posible la expresión de la intersubjetividad, que es premisa y característica determinantes en el altruismo humano.

Capítulo 6.- Intersubjetividad, racionalidad plural y selección de grupo

El sujeto emerge al mundo al integrarse en la intersubjetividad. La intersubjetividad es el tejido de existencia de la subjetividad, el medio de existencia del sujeto, sin la cual él perece. Pero al igual que el individuo no se disuelve ni en la especie, ni en la sociedad, que están en él como él está en ellas, el sujeto no puede disolverse en la intersubjetividad que, sin embargo, le asegura su plenitud (Morin 2004: 82).

... la empatía es la chispa de la preocupación humana por los demás, el pegamento que hace posible la vida social (Hoffman 2000: 3).

No sólo tenemos nuestros propios intereses. También *compartimos* intereses. Y no sólo tenemos nuestros propios motivos, deseos, emociones y talentos. *Compartimos* nuestros motivos, deseos, emociones y talentos. Y esto no sólo consiste en tener los mismos intereses, emociones, etc., al mismo tiempo (sin violar la regla cartesiana de que cada persona tiene su propia emoción)... Somos criaturas tanto sociales como emocionales, para las que el entendimiento mutuo -esto es, el entendimiento de las emociones de unos y otros- es esencial. Sin tal comprensión, no es sólo que cada uno de nosotros vendría a ser un sistema emocional cerrado, rico interiormente pero abstraído ante las necesidades e intereses de los otros. Seríamos algo así como autistas, vacíos, indiferentes incluso ante las emociones más básicas, sujetos a fortalezas y a frustración, pero sin nada del drama y la complejidad que surgen de vivir en un mundo emocional compartido (Solomon 2008: 34).

El odio racial, al igual que el sentimiento místico de comunidad, no son fenómenos naturales innatos en el hombre. Son frutos de ideologías (Mises 2009[1949]: 201).

6.1.- Introducción

En el primer capítulo de este trabajo se concluye en la necesidad de disponer de una teoría del comportamiento humano que incluya la acción motivada en la utilidad de otros. Esto contempla la utilidad para un conjunto de individuos, o sea, un grupo, y de ahí la hipótesis de la acción comprendida, sentida y motivada no por el interés propio, sino por lo que podría llamarse interés o bien común. En el presente capítulo, que es una continuación coherente del anterior, se estudia cómo funciona la mente humana en estos casos, es decir, la psicología de la intersubjetividad desde la perspectiva de la integración psicológica del individuo con el/los individuo/s con que se identifica.

En el apartado 2 se expone la hipótesis del comportamiento individual desde una perspectiva no egocentrada, sino plural o colectiva. Este patrón de comportamiento, que coexiste con el egoísmo psicológico, consiste en actuar de acuerdo a parámetros supra-individuales, es adaptativo y por tanto tiene sentido evolutivo. Su existencia amplía el campo de posibilidades de selección en la especie, y desplaza el análisis hacia el modelo de selección en varios niveles.

En el apartado 3 se presentan argumentos de lo que podría decirse una desmitificación del solipsismo por la constatación del fenómeno de la identidad y de las identificaciones. Existen evidencias de los innatismos subyacentes al instinto de identificación, el cual hace posible la intersubjetividad y el altruismo psicológico.

En el apartado 4 se analiza la neuropsicología del altruismo. Para ello nos apoyamos en diversos estudios experimentales que corroboran su operativa psicológica y fundamentación neurocientífica. Una importante derivación es la afectividad y la cognición desde la perspectiva del nosotros (*we sense*), o perspectiva plural, de la cual se desprenden implicaciones igualmente remarcables desde el punto de vista de la selección grupal a través de conductas de favoritismo inclusivo.

En el apartado 5 se explica este fenómeno, así como la influencia de la ontogénesis cerebral, a través de la asimilación de cultura, en el refuerzo selectivo de la cooperación motivada por procesos neuroendocrinos de afiliación y vinculación.

En el apartado 6 se analizan los procesos selectivos relacionados con la semejanza neural y el fenómeno del acoplamiento neural (*rapport*) de cerebros fenotípicamente semejantes. Los cerebros semejantes comparten una intersubjetividad más rica y extensa, de lo cual se deriva una tendencia a cooperar selectivamente (homofilia), y a practicar comportamientos altruistas inclusivos y discriminatorios, aunque no necesariamente de tipo endofamiliar. Estas tendencias poseen valor de supervivencia debido a los resultados diferenciales de la búsqueda, relación y vinculación con individuos semejantes.

Se concluye el capítulo con una reflexión acerca de las posibilidades que ofrecen la educación y la propagación de replicadores culturales en la estimulación de la benevolencia, y la extensión del comportamiento altruista hacia agregados poblacionales de tamaño y complejidad superior.

6.2.- El comportamiento desde la perspectiva colectiva

Si el sujeto es capaz de identificarse con experiencias ajenas y con experiencias vividas comunitariamente, o comunidades de experiencias, puede por ello actuar sobre la base de una intencionalidad común, que incluye conocimientos, expectativas y deseos de nivel poblacional. Dado que estas actuaciones no conciernen a un solo sujeto, sino que son producto de la intersubjetividad, ello invita a plantear la hipótesis de los procesos de acción plural seguidos por un yo trans-personal. Si estos tienen efectos adaptativos, como es el caso, la hipótesis a dilucidar es la de una inteligencia social o colectiva, que surge de la interacción de inteligencias individuales.

La investigación básica de estos fenómenos tiene una larga tradición académica en la que destaca el pensamiento de Hayek sobre los órdenes espontáneos, así como la teoría de la complejidad, y más cercana en el tiempo la teoría del caos. En la psicología y la ciencia cognitivas contemporáneas existe igualmente un acervo de investigación que muestra la existencia de acción grupal inteligente, y con resultados racionales, impulsada por objetivos y motivaciones, cogniciones, intenciones, identidades y racionalidad colectivas (Malone y Bernstein 2015; Surowiecki 2005). Las mentes, no obstante, no existen pre-acopladas, sino que el acoplamiento es la consecuencia de

procesos de convergencia psicológica a veces largos y complejos. Dicho acoplamiento tiene una causalidad próxima explicada en adaptaciones que operan mayoritariamente en el plano inconsciente, de manera que las decisiones no se toman del todo controlada y deliberadamente, como en la teoría del actor racional, sino que surgen «como si salieran de las entrañas» de un agente dotado de racionalidad ecológica de fondo (De Waal 2011: 23). En estos casos los grupos funcionan no sólo teórica o metafórica, sino experimentalmente, como unidades inteligentes de decisión.

Un estudio extraído de la etología es el de la toma de decisiones grupales en los enjambres de abejas melíferas silvestres (Seeley, *et. al.* 2006). Las abejas desarrollan procesos de decisión colectiva que revelan la inteligencia subyacente a la interacción del grupo. La decisión acerca de cuándo y dónde trasladar la colmena es producto de un procedimiento mediante el cual unas 100 abejas exploradoras de las 10.000 que pueden integrar un enjambre van “proponiendo” alternativas a través de señales expresadas con la danza del ocho. En el proceso, que puede durar hasta 16 horas, se va estableciendo una mayoría estable dentro de esas 100 abejas en torno a una de las alternativas. Cuando se ha alcanzado un quorum determinado (que no un consenso de las propuestas enfrentadas), o elevado índice de saturación de la señal, entonces la decisión se produce, la norma grupal se establece, y el enjambre en conjunto la ejecuta. Este *quorum sensing* es un producto evolutivo para la toma de decisiones grupales, y se da entre alternativas en competencia que existen porque también son el resultado de procesos selectivos epigenéticos que permitieron al grupo sobrevivir en el pasado.

Los estudios llevados a cabo por Frey, *et. al.* (2010a, 2010b, 2010c, 2011) sobre los naufragios de las embarcaciones Titanic y Lusitania, producidos en 1912 y 1915 respectivamente, han esclarecido que en el primero de ellos el pasaje y la tripulación actuaron como si el grupo siguiera una lógica de comportamiento orientada a proteger a los pasajeros más vulnerables, como es el caso de las mujeres y los niños. En el segundo, por el contrario, sobrevivieron principalmente los hombres jóvenes y adultos sanos. Siendo las embarcaciones, el pasaje y la tripulación semejantes, ¿qué puede explicar esta diferencia del comportamiento entre uno y otro percance? En el naufragio del Titanic el colectivo se comportó siguiendo un patrón de comportamiento organizado,

regulado y coherente con las normas sociales (y epigenéticas) de protección de aquellos individuos más débiles, y a la par con superior capacidad biológica de transmitir replicadores a largo plazo. En el hundimiento del Lusitania parece como si el patrón dominante hubiera sido la desordenada lucha por la supervivencia y la victoria de los individuos más fuertes, sin considerar normas sociales ni intereses más allá de los individuales. Dado que toda especie practica comportamientos estereotipados y ordenados en situaciones de emergencia y tragedias colectivas, los resultados finales de ambos sucesos deberían haber sido más parecidos. Una explicación verosímil de la diferencia puede ser que la generación de un quorum requiere tiempo, y que el tiempo necesario para el reconocimiento de la situación, y estimulación de las emociones conducentes a la acción prosocial, pudo ser suficiente en las 2 horas y 45 minutos que transcurrieron desde el impacto con el iceberg hasta el hundimiento del Titanic y, sin embargo, resultó insuficiente en el caso de los sólo 18 minutos que duró el naufragio del Lusitania tras ser atacado por un barco de guerra. En efecto, se ha demostrado experimentalmente que la comunicación más duradera y rica entre individuos genera disposiciones y comportamientos proclives a la empatía y el altruismo, con resultados decantados a un mayor predominio prosocial (Andreoni y Rao 2011; Balliet 2010). En otro orden de cosas, también existen evidencias de que, en las evacuaciones simuladas en situaciones de peligro, los individuos tienden a comportarse más deferentes y dispuestos a ayudar, o sea, más cooperativos, cuando existe una conciencia desarrollada del destino común, de una identidad de grupo, y se dan unas circunstancias menos estresantes en el desarrollo de la acción (Kretz 2011). Obviamente, el reconocimiento de la comunidad de destino e intereses, reconocimiento y comunicación de identidades, etc., conllevan costes de transacción comunicativa y operacional que, en circunstancias de juego de tirada única y escasez de tiempo, decantan el comportamiento hacia el patrón no-cooperativo, como ocurrió en el caso del Lusitania atacado y hundiéndose rápidamente.

La generalidad de animales sociales cuentan con adaptaciones de este tipo, que impelen al individuo a actuar de acuerdo a la norma grupal (quorum) cuando percibe con certidumbre que el ambiente está saturado de la señal que refuerza la expectativa de que una mayoría suficiente acata una determinada norma. Se ha podido comprobar que

los procesos relacionales en que los grupos humanos desarrollan o se percatan de la existencia de elementos de destino común generan procesos de convergencia psicológica conducentes al surgimiento de identidades colectivas, segregación de entramados sociales, acción progrupal y manifestaciones conductuales y verbales de lo que se puede llamar una perspectiva de racionalidad grupal, o espíritu de comunidad (*we perspective, we attitude, we mode*) (Kimbrough, *et. al.* 2008). El comportamiento cooperativo en estos casos puede ser explicado teóricamente mediante la hipótesis de la identificación grupal y una de sus consecuencias, que es el razonamiento de equipo:

... el razonamiento de equipo (*team reasoning*) difiere y tiene mayor poder explicativo que asumir el objetivo del grupo y sobre ello razonar de la manera individualista estándar. La particularidad de esta capacidad explicativa incrementada estriba en que los individuos que razonan como equipo adoptan un punto de vista común, más amplio (que el individual). (Bacharach 1999: 144; paréntesis añadidos).

La investigación neurocientífica ha localizado la que puede ser la base de los dispositivos y procesos neurológicos, o sea la causalidad próxima, de estas adaptaciones en los primates (Reardon 2014). Existe evidencia de un dispositivo neurológico humano adaptado para la cognición social, que sólo se activa de forma complementaria a otros dispositivos generales de inferencia, y que lo hace cuando es preciso considerar hipótesis y razonamientos relacionales (Mason, *et. al.* 2010). La actividad neurológica se activa o se incrementa en determinadas áreas cuya función era hasta hace poco tiempo desconocida (corteza medial parietal y dorsomedial prefrontal, además de las ya conocidas de la corteza frontal inferior, la temporal superior, y el giro fusiforme) cuando se perciben o participa en procesos relacionales (Iacoboni, *et. al.* 2004). De este modo, el cerebro tiene un funcionamiento diferente cuando el espectador observa imágenes de interacción, de cuando observa acciones solitarias cotidianas o cuando se halla en estado de reposo.

Las adaptaciones que capacitan para la toma de decisiones grupales se basan en la capacidad básica de la diada de individuos que actúan de acuerdo a un propósito que trasciende sus intereses individuales. Esto incluye tanto la cooperación para el logro de objetivos transversales como el puro altruismo del uno al otro. En ambos casos se manifiesta incumbencia, o interés por utilidades que trascienden el propio interés individual estrechamente entendido:

El donante incluso puede considerar los activos del receptor como parte de su propia estructura de activos, de forma que el regalo se parece más a una distribución de activos entre las partes de una misma organización. (Boulding 1976: 15).

Estos procesos decisorios requieren la armonización y sincronización de aspectos cognitivos, emotivos e intencionales determinantes de la acción.

En el ámbito de discusión de los procesos de causación última del comportamiento grupal Harsanyi y Selten especifican, en su libro *General theory of equilibrium selection games* (1988), que el principio de coordinación no puede derivarse de la racionalidad individual, sino que implica necesariamente la existencia de un modo de racionalidad colectiva, nudo gordiano del patrón psicológico cooperativo de la especie, que los autores explican en los siguientes términos:

(La racionalidad colectiva)... se basa en el supuesto de que en la ausencia de alguna razón especial en contra, los jugadores racionales elegirán un punto de equilibrio que les reporte a todos ellos ganancias superiores, en lugar de uno que les reporte ganancias inferiores. O sea, se basa en el supuesto de que los individuos racionales cooperarán en la persecución de sus intereses comunes si las condiciones permiten que lo hagan. (Citado en Schmidt 2007: 169; paréntesis añadido).

Consecuentemente, Schmid (2005a: 62) apostila que: «No todo lo que nos motiva procede de nuestra colección subjetiva de motivos. Algunos son intersubjetivos». En esta misma línea, Tuomela (2000: 12) añade que:

Actuar conjuntamente involucra compartir la acción en un sentido plural (*we sense*) (“estamos haciéndolo juntos”). Debe existir unidad o proximidad en la cooperación, al menos en el sentido de que los participantes tienen la intención de actuar conjuntamente y basan su acción en la expectativa de que otros lo harán así al menos con alguna probabilidad. Los agentes han de compartir mínimamente la acción conjunta en la que cooperan.

El solipsismo parece no corresponderse con estos presupuestos. Pasemos a esbozar una breve reflexión crítica sobre esta influyente teoría filosófica.

6.3.- Desmitificación del solipsismo

“Otras mentes” es un tema filosófico clásico, uno de cuyos problemas a tratar es el que plantea la negación de la posibilidad de tener acceso a la vida mental de otros individuos. Esta hipótesis ha sido denominada *perspectiva egocéntrica*, y es un enfoque influyente tanto en filosofía de la mente, como en las ciencias de la decisión y del comportamiento (García Suárez 1976: 20 y ss.) En esencia se trata del problema del solipsismo, epistemológico y motivacional, un supuesto teórico según el cual no es posible conocer la interioridad ajena, y/o dicha interioridad no resulta intrínsecamente motivadora del comportamiento de quien la pueda conocer³⁵⁶. Esta *perspectiva egocéntrica*, que representa la visión cartesiano-lockeana de la vida mental del sujeto, desemboca en un individualismo atomista dentro del cual no es posible la cooperación, y el altruismo, que no tendría una lógica explicativa salvo el azar o el capricho en la conducta, no podría contar con posibilidades de asignación eficiente. Ello se debe a que la concepción solipsista de la mente es la de «un solitario lugar de representaciones», desconectado de las vivencias de otros yoes y, por lo tanto, incapacitado para establecer identidades grupales, e intenciones y acciones comunes (Schmid 2003: 205). El problema de las otras mentes es una reflexión escéptica sobre el contenido y alcance del yo. No obstante, tras un largo debate intelectual, el solipsismo en su negación de la intersubjetividad puede considerarse hoy día una teoría filosófica y una ideología refutadas por la propia filosofía y por la ciencia³⁵⁷. Como afirma Edgar Morin (2004: 87): «El sujeto no está solo porque el Otro y el Nosotros están en él.»

En efecto, todo estudio de la intersubjetividad parte de que es un hecho constatado que la experiencia humana admite la suavización, si bien selectiva y limitada, de lo que podemos llamar fronteras del sujeto. Ya Freud asimiló la experiencia universal del enamoramiento, por ejemplo, a una de disipación del yo solipsístico. Sin

³⁵⁶ Muy esquemática y simplificada, el solipsismo se trata de una doctrina que afirma “sólo existo yo” (Kenny 2005: 462). Dowding y King (1995: 14) aluden a una visión egocentrada del individuo en la que «el mundo se ve exclusivamente desde la perspectiva del sujeto que toma las decisiones y no desde un punto de vista social». La principal consecuencia de esta visión extrema del egoísmo psicológico consiste según Arrow (1974: 21) en que: «por definición, (se) puede ignorar al resto del mundo» (paréntesis añadido).

³⁵⁷ Un análisis crítico acerca del solipsismo realizado desde la filosofía de la ciencia es el de Deutsch (1999). Un clásico estudio crítico similar, dentro de la filosofía analítica, es el de Ryle (1967[1949]). El descubrimiento de las neuronas espejo ha permitido a la neurociencia refutar experimentalmente la teoría filosófica solipsista (Iacoboni 2009; Rizzolatti y Sinigaglia 2006).

embargo, tal disipación de la conciencia egocentrada puede ocurrir con la misma intensidad motivacional en otras experiencias y circunstancias habituales. Así, Feinberg (2001: 149) subraya que: «Los márgenes del “yo” no son completamente fijos, sino que se modifican durante la vida del individuo». En efecto, como ya recogiera Adam Smith en diversas partes de su obra, adelantándose en casi dos siglos y medio a los posteriores descubrimientos neurocientíficos, los márgenes del yo se ven involuntariamente afectados por la simpatía con los demás, uno de cuyos estímulos acrecentadores es la contemplación del sufrimiento ajeno y la necesidad de ayuda (Smith 1997[1759]: I, i). Para Smith, y también para Darwin, que se inspiró en los trabajos de aquél, la simpatía es un instinto, y como tal no se trata de una disposición que surja de la decisión deliberada y racional. Smith tiene claro que sentimos una innata y poderosa curiosidad instintiva por la vida mental de los otros, y habla de ello como:

Esta pasión por descubrir los sentimientos reales de los demás naturalmente tan poderosa que a menudo degenera en una curiosidad molesta e impertinente. (Smith 1997[1759]: VII, iv. p. 589).

Asimismo, Smith especifica que la atracción simpatizante es más intensa cuando se produce una conexión mental amplia y profunda, y esta conexión (que en psicología se denomina actualmente *rapport*), se da con mayor facilidad entre mentes que tienen puntos en común, y que por lo tanto se expresan en cerebros filogenética y/u ontogenéticamente similares en algún aspecto relevante (Smith 1997[1759]: VII, iv. p. 588). Los cerebros simpatizantes manifiestan analogías (e incluso homologías) en los mapas y modelos emotivo-cognitivos, que son el sustrato neurofisiológico de las identificaciones e identidades personales.

El conjunto de individuos neurológicamente semejantes en una determinada estructura o modelo encefálico son una *forma de vida* caracterizada por transportar y replicar una determinada representación del mundo, un *Weltbild*, utilizando una expresión filosófica (Kenny 2005: 466). Este fenómeno, que puede tener un sustrato tanto genético como cultural, permite la superior conexión simpatética en la interacción, y esa es una causa explicativa de la preferencia por la semejanza (Gallese y Goldman 1998: 498).

La refutación del solipsismo tiene una gran trascendencia en las ciencias sociales, dado que otro de los debates filosóficos clásicos, el que alude al egoísmo y a la concepción estrecha del interés propio, resulta igualmente refutado. Con esto cobran sentido las matizaciones al enfoque de la que Hirshleifer (1985: 55) denominó «simple estrategia pragmática del hombre egoísta». El interés propio, lo incumbente para el sujeto, parte de una preocupación activa por la familia y los amigos, principalmente, aunque es extensible, parafraseando a Hayek, a «todo aquello por lo cual la gente de hecho se preocupa» (Hayek 1948: 13). Como vamos a explicar, las preocupaciones de la gente por los demás, y de ello sus comportamientos prosociales, están positivamente correlacionadas con la variable semejanza. Por esta razón la pregunta clave en los procesos de identidad e identificación no es tanto «¿quién soy yo?», sino «¿a quién me parezco?»

6.3.1.- «¿Quién es como yo?»

El conocimiento sobre «¿quién es como yo?» es una inquietud primaria en la especie humana, estudiada en los campos de la antropología social y cultural (Gil-White 2001) y de la neurociencia (Kaplan, *et. al.* 2007). En las relaciones empíricas y en los sistemas ideológicos la idiosincrasia del otro no es en modo alguno indiferente, y los experimentos de neuroimagen muestran que la atribución y pertenencia a las identidades sociales influye en la respuesta neurológica determinante de las actitudes de atracción o rechazo. La motivación altruista viene explicada por la identificación, y esta por la activación de procesos neuropsicológicos asociados a la afiliación y afinidad.

En el capítulo 5 explicamos que el yo teorizado por Hallowell es un yo social y empírico, identificado con un conjunto coherente de normas culturales. Además es un sujeto evaluador de concordancias y discordancias entre el comportamiento propio, la identidad personal y la proyección social de esa identidad. La coherencia entre estas tres facetas del sujeto es buscada activamente de manera autónoma, y las discordancias producen un malestar psíquico relacionado con los fenómenos de la disonancia cognitiva y la crisis de identidad. Estas capacidades, que contribuyen a estabilizar la identidad personal y social del individuo, son facultades mentales superiores, características de un animal profundamente cultural, como el humano, y que capacitan

para una interacción social regulada por normas con sentido moral, dentro de las que cabe el altruismo cultural, o sea, un altruismo suprafamiliar³⁵⁸.

Hallowell asoció esta modalidad de altruismo, que él llamó *moral*, a la capacidad propiamente humana de comportarse según reglas culturales, lo cual ya es, en sí mismo, una disposición prosocial instintiva y coherente con los modelos de selección grupal³⁵⁹. Esta capacidad requiere una capacidad psicológica previa de auto-evaluación y evaluación de los otros yoes de acuerdo a las normas (epigenéticas y culturales) que regulan la vida social (Turner 2000: 139-140). La interiorización y el seguimiento de un sistema de normas culturales refuerza el sentido de ser algo, de tener una identidad, y existe una disposición psicológica innata en pos de la estabilidad identitaria sobre la base de la auto-evaluación.

Otra dimensión de la identidad y la identificación, en este caso no con normas abstractas o marcadores simbólicos, sino con individuos concretos, es la que guarda relación con las capacidades humanas relativas a la intersubjetividad. La intersubjetividad parte del umbral de una díada de individuos, pero puede proyectarse por inferencia a una multitud abstracta, dado que los humanos extrapolan la psicología de la primera persona a una racionalidad grupal. Ello supone que los individuos conciben a los grupos como si de individuos compuestos se trataran. Esta característica transforma el sentido de la individualidad y saca al sujeto de su aislamiento egocéntrico, abriendo una dimensión social y amplificada de la función de utilidad, de las metas que persigue el individuo, y de los procedimientos para conseguirlas. Téngase en cuenta que, como indican Tyler y Blader (2001: 211 y 213), la identificación social de los individuos «pone en marcha una dinámica psicológica que transforma el sentido interno de la individualidad», y este proceso se lleva a cabo a través de: (a) la integración psicológica del yo y el grupo con que se identifica (conciencia y motivación desde el “nosotros”); y (b) la combinación del sentido existencial (*sense of self*) del individuo y

³⁵⁸ Las facultades mentales humanas, explicativas de capacidades específicas como el altruismo no consanguíneo, obedecen a un cerebro anatómicamente distinto del de otros animales, y concretamente de los primates sociales. Darwin se percató de ello, y atribuyó las facultades morales humanas no a una cualidad única en el reino animal, sino a una cuestión de grado y complejidad. Véase (Darwin 1972[1871]: I, 4. p. 95).

³⁵⁹ Hallowell pudo inspirarse en la obra previa de Edward Westermarck, quien plantea una hipótesis muy parecida sobre la capacidad de identificación y la consecuente motivación basada en emociones sociales auto-reguladas (véase De Waal 2006: 19-20).

el grupo, o sea, la conciencia de un destino común. Estos procesos posibilitan la acción de grupo y la selección en varios niveles, y resultan explicativos del altruismo humano sobre la base de una racionalidad supraindividual, de grupo y de especie. ¿Qué tipo de neuropsicología posibilita tal dimensión del altruismo?

6.4.- La neuropsicología intersubjetiva del altruismo

Si se considera el comportamiento altruista sólo como un asunto empírico-positivo, definido únicamente como la entrega de recursos o realización de comportamientos de ayuda con un coste para quien los practica, y desligado de su contenido psicológico, es posible desenfocar teórica y analíticamente el fenómeno³⁶⁰. Un comportamiento puede ser empíricamente altruista pero intencionalmente egoísta (Andreoni 1989, y 1990). La especie *Homo sapiens* dispone de una psicología social especializada en descifrar la intencionalidad de los comportamientos (deseos, objetivos y planes de acción), porque la selección del rasgo altruismo requiere, en el caso humano, la clarificación de las intenciones³⁶¹. La posibilidad de un altruismo utilitario, camuflado de una aparente preocupación por el objeto receptor de la ayuda, puede desviar el comportamiento de la eficacia inclusiva y situarlo realmente en coordenadas de selección de vehículos, no de replicadores. Y esto es así porque

... el espíritu de cálculo que motiva al tramposo sutil (...) puede en el futuro llevarlo a hacer trampas cuando las condiciones parezcan más ventajosas (...) Existe amplia evidencia para soportar el argumento de que los humanos responden a actos altruistas de acuerdo a su percepción de los motivos del altruista. Se tiende a responder más altruistamente cuando se percibe al otro actuando de manera “genuina”, esto es, realizando voluntariamente un acto cooperativo como un fin en sí mismo, sin esperar ganancia (...) Específicamente, la selección puede favorecer el aprender de experiencias altruistas y tramposas de otros, ayudando a otros a castigar tramposos, construyendo una diversidad de sistemas de intercambio, y formulando reglas para regular dichos sistemas. (Trivers 1978[1971]: 216-220).

³⁶⁰ Por ejemplo, en la teoría económica del altruismo se opta explícitamente por un enfoque conductista de preferencias reveladas, en el que resulta irrelevante la motivación y la intención a efectos teórico-analíticos. Véanse a este respecto Becker (1997[1976a]: 337), y Becker (1981).

³⁶¹ El sistema de neuronas espejo desempeña esta función psicológica, por la que el individuo es capaz de comprender y evaluar las intenciones ajenas de manera automática, a través de la simulación mental con base en la información aferente captada por los sentidos, y el relleno de *qualia* informativos mediante la extrapolación. Para que esta operación funcione correctamente, el cerebro dispone de una red neural especializada en diferenciar la identidad del individuo que está llevando a cabo las acciones. La simulación y comprensión (cognitiva y emotiva) de las intenciones se da independientemente de la identidad del agente observado, si bien las semejanzas filogenéticas y ontogenéticas de los cerebros favorecen una operativa más precisa y resultados más acertados. Véanse al respecto Adolphs, *et. al.* (1998); Stone (2000); Urgesi, *et. al.* (2007); Preston y De Waal (2002b).

Existen pruebas experimentales de que el cerebro decodifica intenciones y estados emotivos a través de procesos funcionales endógenos y especializados. La base neurológica de esta adaptación para la interacción social ha sido descubierta recientemente en primates no humanos (Horoush y Williams 2015), así como en humanos (Koster-Hale, *et. al.* 2013; Ma, *et. al.* 2011; Engell, *et. al.* 2007; Todorov, *et. al.* 2007). La operativa se despliega a través de las capacidades para la intersubjetividad, y especialmente las de teoría de la mente³⁶². Asimismo, los organismos dotados de teoría de la mente infieren esos estados mentales en otros a través de la observación de signos externos que señalizan información, lo cual permite establecer hipótesis y teorías sobre la intención y actitud ajenas adaptadas a cada circunstancia (Heyes 1998: 101-102). De este modo, los individuos no incapacitados para estas operaciones, como sería el caso de las personas aquejadas de autismo y niños pequeños con edades inferiores a los 3-4 años, pueden diferenciar entre lo accidental y lo intencional cuando realizan juicios morales y evaluaciones sobre los comportamientos observados en otros. La atribución de creencias, intenciones y actitudes es la base epigenética del juicio moral, y tiene una influencia directa en la motivación y en el comportamiento social (Young, *et. al.* 2007; Snyder y Stukas 1998).

Esta capacidad epigenética puede tener un papel relevante, además, en la explicación de la filogénesis de los circuitos neurales flexibles que producen la acción motivada por contagio emocional (que puede ser empíricamente altruista, aunque psicológicamente es egoísta), como dispositivo diferenciado del que explica la acción motivada por simpatía (que psicológicamente es altruista) (Preston y De Waal 2002b: 13-14). En efecto, una gama predominante de los comportamientos de altruismo intencional se explica porque el individuo realiza comportamientos de ayuda no por aliviar la tensión emocional que le produce el contagio de las emociones y sentimientos de malestar observados en el otro (*empathic distress*), sino porque resulta motivado por ese conocimiento (*empathic concern*), sabiendo discernir con claridad a quién pertenecen esas emociones y sentimientos, y también a quién incumben directa o indirectamente (Vignemont 2006). El altruismo por preocupación empática es un comportamiento consciente de preocupación activa por las necesidades ajenas, en el que el foco de

³⁶² La teoría de la mente consiste en la suposición innata de que los organismos tienen estados mentales que causan sus intenciones y comportamientos.

atención que desencadena la motivación y el compromiso prosociales no apunta al individuo observador, sino al individuo observado: «Si no se es capaz de sentir la diferencia entre el propio dolor de uno y el dolor de otro que sufre, no existe base para la simpatía» (Wispé 1991: 158). Los comportamientos empíricamente altruistas de los primates no humanos (con excepción, y de forma muy primaria, de alguna especie, como es el caso de los chimpancés) son motivados principalmente por el alivio del estrés empático, e igualmente muchos de los realizados por los individuos humanos (p.ej. dar una limosna porque resulta incómoda la presencia del mendigo suplicante, y no por la simpatía con su situación de necesidad)³⁶³. Así, lo específico del comportamiento altruista humano no es actuar por contagio emocional, sino ayudar al otro sintiéndolo como una extensión del sí mismo, aunque con la consciencia de que se trata de un yo diferente y autónomo³⁶⁴.

Para la neurociencia actual la empatía consiste en la capacidad de coparticipar en los sentimientos de otros individuos. La identificación empática permite actuar motivado por emociones y sentimientos de otro individuo. Existe evidencia empírica de la coactivación isomórfica, simultánea y sincronizada de áreas cerebrales durante la experiencia empática, lo que ha llevado a plantear la hipótesis de la experiencia en red *-shared network hypothesis-* (Vignemont y Singer 2006)³⁶⁵.

En la ontogénesis del espécimen humano, los bebés pequeños con edades inferiores a los 3-3,5 años todavía no disponen de un cerebro suficientemente desarrollado para la intersubjetividad y la cognición social, ni a través de la empatía, ni tampoco de la teoría de la mente. Ello se debe a que su corteza prefrontal aún no se ha desarrollado lo necesario. Ésta es el área anatómica del cerebro especializada en la

³⁶³ Explica De Waal (1997: 75 y 93) que la mayoría de los comportamientos de ayuda que se realizan en el reino animal están motivados por un intento de aliviar la angustia asociada al contagio de emociones desagradables, como el miedo, o la conciencia de emociones asociadas a la experiencia de dolor o sufrimiento en general. La acción de ayuda por contagio emocional es psicológicamente egoísta, dado que está motivada por la aversión a sentimientos desagradables reflejados en el observador.

³⁶⁴ Khalil (2004: 116) se refiere a este fenómeno mental de amplificación de la frontera de objetos de incumbencia, con la denominación de “yo expandido” (*expanded self*).

³⁶⁵ No obstante, la coactivación isomórfica de neuronas singulares, o de redes de neuronas diferenciadas, no deja de ser una hipótesis de trabajo en fase de experimentación, dado que hoy día no existe una tecnología suficientemente desarrollada para detectar y medir actividad neural a esos niveles de desagregación (Singer y Lamm 2009: 87).

lectura de otras mentes a través de la inferencia analógica (Preston y De Waal 2002b: 14). Aunque los niños pequeños ya disponen de un proto-yo desde los 1,5-2 años (utilizan pronombres personales, se reconocen a sí mismos ante un espejo, manifiestan emociones como la vergüenza y conductas de engaño), sin embargo sólo a partir de los tres años surge un nivel de conciencia más avanzado, que es la conciencia reflexiva, o subjetivación³⁶⁶. Como precisa Austin (2006: 13), es a esa edad cuando los niños se encaminan a la intersubjetividad, y ya entre los 4 y 5 años su cerebro está capacitado para una cada vez más compleja mentalización de las relaciones con otros yoes³⁶⁷. Esto quiere decir que el niño/a a esa edad comienza a realizar inferencias y proyecciones de los contenidos y estados mentales de otros sujetos sobre la base de su propia mente. La capacidad de leer otras mentes aparece con anterioridad a la concepción subjetiva de la categoría del tiempo y su repercusión en la comprensión de la identidad personal. Es decir, el cerebro infantil está capacitado antes para una intersubjetividad muy elemental que para concebirse a sí mismo como entidad social y cultural que permanece en el tiempo³⁶⁸. Las bases neuropsicológicas para comprender los rasgos definitorios del autoconcepto histórico-cultural y su estabilidad surgen con posterioridad a las

³⁶⁶ Más o menos a los 3 años los niños y niñas pequeños ya disponen de un dispositivo neurológico incipiente que les permite conocer qué le está ocurriendo subjetivamente a otros individuos, así como comprender vicariamente las emociones y sentimientos de aquéllos. Esta disposición empática (cognitiva y afectiva) ha sido estudiada con una población de 122 gemelos por Knafo, *et. al.* (2009), quienes han encontrado que el grado de comprensión empática está explicado por factores genéticos en una horquilla entre el 19% y el 44%. El resto se explica por factores ambientales y error estadístico.

³⁶⁷ La mentalización consiste en la operación de comprender contextualmente el sentido y significado de las acciones propias y ajenas, a través de representaciones mentales basadas en entidades inobservables, como son las creencias e intenciones (Tooby y Cosmides 2007: 43). La principal función de esta adaptación neuropsicológica, localizada físicamente en el área de conjunción temporo-parietal derecha con otras redes adyacentes, es llevar a cabo inferencias acerca de lo que otros individuos piensan sobre los pensamientos propios (Saxe y Kanwhiser 2003; Saxe y Wexler 2005). Aunque se trata de una habilidad cognitiva compleja, no por ello aparece sólo en individuos adultos. Ya en la primera infancia comienza a manifestarse una proto-mentalización que produce representaciones elementales de causalidad, aunque de contenido básicamente emocional y semiconsciente (Saxe y Carey 2006). Sólo después de los cuatro años de edad los niños pequeños empiezan a realizar inferencias conscientes sobre los estados mentales ajenos (Gweon, *et. al.*, 2012; Singer 2006: 856).

³⁶⁸ Según Lewis (2005: 707) el orden de los pasos ontogenéticos que conducen a una vida mental apta para la cognición social son: «(1) Yo sé, (2) yo sé que sé, (3) yo sé que tú sabes, y (4) yo sé que tú sabes que yo sé». De este modo, no es posible cooperar socialmente de acuerdo a intenciones compartidas sin un desarrollo previo de la autoconciencia. No obstante, el bebé sí dispone de un proto-yo con un nivel muy básico de conciencia desde que nace. Ese proto-yo es susceptible de la intersubjetividad primaria de la que estamos hablando. Esta intersubjetividad elemental que tienen los bebés de hasta aproximadamente 2 años, consiste en la afectación emocional de otros a través del contagio de emociones y una primaria comprensión refleja de lo que les ocurre, similar, homológicamente hablando, a la intersubjetividad de otras especies de grandes primates, como los chimpancés. Véanse Suddendorf y Whiten (2001) y Gweon, *et. al.* (2012).

capacidades de, en primer lugar, resultar afectado (contagiado) por las emociones ajenas (un aspecto del desarrollo psicológico que se manifiesta desde el nacimiento), y en segundo lugar, la capacidad de empatizar, o sea, de situarse no en las emociones ajenas y sentir las acriticamente como propias, sino de posicionarse en la situación del otro, y comprender sus emociones sin asumirlas como propias, aunque puedan resultar incumbentes³⁶⁹. Así pues, en la ontogénesis del cerebro humano se desarrolla antes la capacidad de motivación para realizar proto-comportamientos altruistas, que la de concebir una identidad personal histórica, social y cultural, muy relacionada, ésta última, con la aptitud para hacer inferencias que requieren una teoría de la mente adaptada a coordenadas culturales específicas (Singer 2006: 860-861). Esto explica que la intersubjetividad subyacente a los comportamientos prosociales es un proceso automático, que no requiere un previo desarrollo de complejas competencias cognitivas ni de raciocinio, ni tampoco una auto-adscripción a categorías culturales de identidad (Iaboboni 2009: 152-153). La neuropsicología del altruismo es innata, y por lo tanto un rasgo evolutivo independiente de la naturaleza humana, que también se encuentra incipiente en el etograma de otras especies de primates (Byrne y Whiten 1988).

Comentemos brevemente a este respecto los resultados de algunos experimentos que pueden ayudar a comprender mejor el funcionamiento del cerebro en los procesos que desencadenan el comportamiento altruista humano:

- a) Cuando las personas observan la imagen de un rostro que expresa una tristeza no simulada, sus pupilas se contraen automáticamente del mismo modo que están contraídas las pupilas de los ojos que aparecen en las imágenes que observan. Esto indica que la tristeza se ha contagiado al observador, y que ambos cerebros están “acoplados” de manera que el observador puede comprender y sentir las emociones del otro, y actuar en consecuencia (Harrison, *et. al.* 2006).

³⁶⁹ Como se extrae de las investigaciones de Singer (2006), Saxe y Kanwisher (2003), y Warneken y Tomasello (2006), las habilidades de mentalización (comprender los pensamientos y otros contenidos mentales ajenos) y empatía (compartir las emociones ajenas, conocerlas y sentir las vicariamente) siguen dos trayectorias ontogenéticas diferenciadas que implican estructuras neurales distintas. La capacidad de empatizar surge antes del lenguaje y la capacidad de mentalización, porque la primera se lleva a cabo en estructuras del sistema límbico que se desarrollan precozmente, mientras que la segunda depende de un desarrollo que es posterior, de las estructuras pre-frontales y áreas específicas de los lóbulos temporales, como el área de conjunción temporo-parietal.

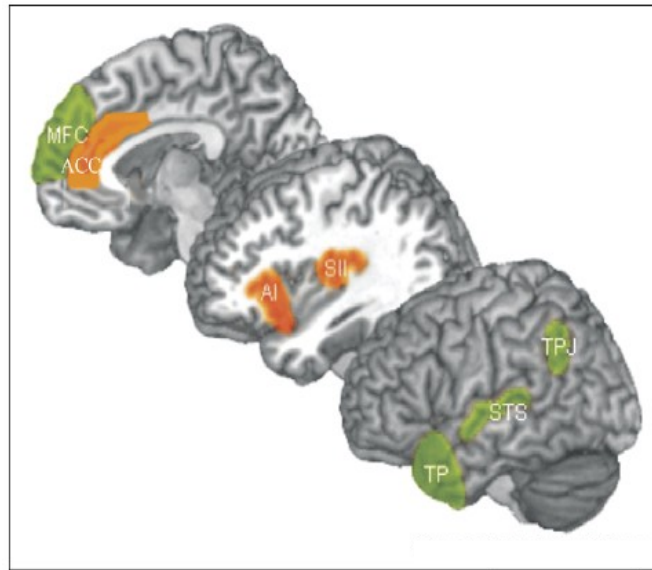
- b) Si se muestran imágenes de personas a quienes los participantes consideran cooperadores o tramposos respectivamente, sus cerebros no procesan de igual forma esa información: las imágenes de cooperadores activan la misma red neural del sistema de recompensa que resulta activada ante el olor de una apetitosa comida o la imagen de una persona sexualmente atractiva, mientras que las imágenes de rostros de tramposos activan la red neural que opera en la producción de emociones de repulsión y desagrado, como olores pestilentes o imágenes de alimentos podridos (Singer, *et. al.* 2004b). Este resultado informa de que la cooperación social se considera en sí misma compensatoria al margen de los potenciales resultados pecuniarios, o de eficacia personal, derivados de la interacción cooperativa³⁷⁰.
- c) Cuando se somete a los participantes al juego del dilema del prisionero secuencial en una sola tirada, los jugadores que cooperan recíprocamente registran una activación intensa del sistema dopaminérgico de recompensa, algo que sólo ocurre cuando cooperan ambos, y que no se da en las restantes posibilidades de actuación en el juego. Esto resulta así a pesar de que no cooperar puede aportar ganancias superiores a sí hacerlo (Rilling, *et. al.* 2004a). La cooperación recíproca lleva asociada una recompensa psicológica en sí misma, que permite superar la atracción utilitaria de la estrategia dominante. Además, la experiencia de cooperación recíproca refuerza la acción de búsqueda de cooperadores, así como el rechazo de no cooperadores en interacciones sociales (en el experimento los participantes responden sistemáticamente con defección a quien ejecuta estrategia dominante la acción previa) y, lo que resulta más sorprendente, los circuitos dopaminérgicos de recompensa se activan ante la cooperación recíproca con humanos, no así con ingenios artificiales, como el caso de un programa informático (Rilling, *et. al.* 2002, y 2004a).

³⁷⁰ En el lenguaje de la teoría de juegos esto quiere decir que la identidad del co-jugador no es neutral, sino que influye en los pagos subjetivos tal y como los computa (consciente o inconscientemente) el agente en su toma de decisiones. Como explican Akerlof y Kranton (2000: 748), la identidad, entre otros efectos, cambia la estructura de las recompensas derivadas de las propias acciones y, asimismo, cambia las recompensas derivadas de las acciones realizadas por otros (es decir, se asocia a efectos externos). Un análisis empírico de estos aspectos realizado con técnicas de neuroimagen puede encontrarse en Fehr y Camerer (2007).

- d) Obviamente, la toma de conciencia de que se está ante un cooperador humano resulta gratificante en sí misma, lo cual vuelve a traer a colación la importancia de los marcadores que señalizan eficazmente los rasgos de la humanidad de lo humano, puesto que esa humanidad reconoce la humanidad del otro, puede identificarse con ella, y esto motiva a una respuesta cooperativa (Fiske 2009). En dilemas sociales sometidos a examen por Rilling, *et. al.* (2004a: 1.700-1.701) mediante técnicas de neuroimagen, se aprecia en los cerebros de los participantes una intensa activación de redes neurales relacionadas con la lectura de otras mentes e identificación empática cuando se coopera con humanos, y no así, o en un grado mucho menor, cuando se interacciona con un programa de ordenador³⁷¹. Aún de este modo, los jugadores que llevan a cabo interacciones de cooperación sabiendo que lo hacen con un programa informático, actúan activando las redes neurales mencionadas, como si estuvieran atribuyendo (errónea o exadaptativamente) una naturaleza humana a su contrincante informático.
- e) En experimentos en que se muestran imágenes de expresiones emocionales de miedo, dolor, sufrimiento y tristeza, en fracciones de segundo resulta activado el sistema neuroendocrino de alerta del observador tras la señal emitida por sus amígdalas. Tal proceso neural se activa incluso cuando las imágenes son proyectadas a una velocidad tan rápida que los participantes no advierten conscientemente que las fotos contienen la expresión de esos estados emotivos (Vuilleumier, *et. al.* 2002; Dimberg, *et. al.* 2000). Estos datos sugieren que observar la expresión de emociones en otros, y experimentar directamente tales emociones en primera persona, activa sistemas neurológicos isomórficos que representan estados mentales semejantes y compartidos (véase figura 6.1.)

³⁷¹ Las adaptaciones neuropsicológicas relativas a la teoría de la mente sólo producen actividad ante la presunción de interacción con un sujeto, es decir, con un organismo que tiene estados mentales, y no así (o en una medida mucho más reducida) con objetos no humanos (Saxe y Kanwisher 2003).

Figura 6.1.- Áreas neurales activas en la intersubjetividad



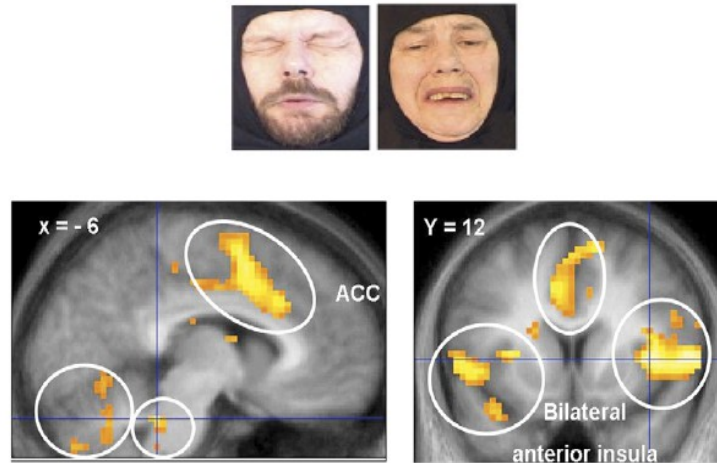
Esquema de las regiones cerebrales involucradas en la comprensión intersubjetiva sobre la base de la teoría de la mente (en color verde) y empatía (color naranja). Esta última representa respuestas isomórficas co-activadas ante la percepción del dolor, asco, gusto y tacto en otros. MPC (Córtex Prefrontal Medial); ACC (Córtex Cingulado Anterior); TP (Polos Temporales); STS (Surco Temporal Superior); TPF (Conjunción Temporo-Parietal); AI (Ínsula Anterior); SII (Córtex Somatosensorial Secundario).

Fuente: Hein y Singer (2008: 154).

- f) Cuando se percibe directamente, o se imagina, el dolor o sufrimiento psíquico de otra persona, se co-activan automáticamente procesos neurales que producen la representación de una experiencia psicológica similar a la que está experimentando quien sufre esa situación desagradable en primera persona (empatía cognitiva). Esto no conlleva la co-activación de todo el entramado neurológico de la experiencia de dolor. El observador no siente la localización ni la intensidad del *dolor físico*, porque en su cerebro entran en funcionamiento los circuitos corticales somatosensoriales primario y secundario. Por el contrario, sí conoce la naturaleza del *dolor emocional* que está sintiendo el otro (su dimensión afectiva), pues en su cerebro se activan tanto la red de procesamiento del dolor (*pain matrix*), que incluye circuitos especializados en la simulación emocional, como los circuitos que operan en funciones de mentalización (Bruneau, *et. al.* 2012; Singer, *et. al.* 2004a, y 2004b). Esta experiencia afectivo-

cognitiva puede desencadenar procesos motivacionales de acciones altruistas (empatía afectiva, o simpatía), si bien la intensidad del co-sentimiento y el resultado final de la vivencia empática dependen de variables y circunstancias que intervienen en la neuromodulación del proceso cognitivo-emotivo-operativo (Hein y Singer 2008; Vignemont y Singer 2006). (Véase figura 6.2.)

Figura 6.2.- Neurología de la experiencia empática del dolor ajeno



Resultados de experimentos con técnicas de neuroimagen sobre estimulaciones neurales de empatía por dolor. Las regiones neurales activadas simultáneamente en el experimento, en el caso de la experiencia de dolor en primera persona y el caso de experiencia vicaria de dolor en otra persona (empatía por dolor): córtex cingulado anterior (ACC) e ínsula bilateral anterior (BAI). La observación de rostros con expresiones de dolor activa zonas del cerebro relacionadas con procesos de empatía, que coinciden con las zonas activadas en la experiencia de dolor producida en primera persona por estimulaciones dolorosas.

Fuente: Hein y Singer (2008: 155).



La actividad neural es simultánea en dos cerebros (del observador simpatizante y de quien tiene la experiencia en primera persona) en la ínsula anterior (AI) y el córtex cingulado anterior (ACC). El estudio de las pruebas del experimento de neuroimagen muestra respuestas neurales isomórficas producidas en los cerebros de individuos que ven fotografías de una situación potencialmente dolorosa (posibilidad de un corte en la mano utilizando un cuchillo de cocina), expresiones faciales de dolor, y una aguja clavándose en una mano humana.

Fuente: Vignemont y Singer (2006: 436).

- g) Si se realizan acciones de cooperación y altruismo que contribuyen con éxito a mejorar la situación vital de otros individuos, surge espontáneamente una

sensación de bienestar que no formaba parte del propósito de la acción, y que en la terminología popular se asocia al concepto de alegría empática (Dunn, *et. al.* 2008; Aknin, *et. al.* 2013). Esto invita a defender la hipótesis de un sistema congénito de recompensa de acciones conducentes a las relaciones prosociales, y la motivación por la supervivencia de los rasgos de humanidad. Así, diversos estudios de neuroimagen han encontrado relaciones entre el altruismo y la activación de los circuitos neurales paralímbicos relacionados con la producción de oxitocina, dopamina y vasopresina por el hipotálamo, sustancias hormonales y neuromoduladoras asociadas a las conductas de confianza, vínculo personal, afecto amoroso y cooperación, cuya presencia en el organismo está asociada a sensaciones de armonía y bienestar (véanse Moll, *et. al.* 2006; Harbaugh, *et. al.* 2007; Tricomi, *et. al.* 2010).

Estos experimentos demuestran que existen procesos neurológicos instintivos subyacentes a los comportamientos de cooperación y altruismo, que se activan automáticamente, y que generan la acción prosocial sin necesidad de una voluntad consciente y premeditada de cumplir con una obligación moral, eludir un castigo, o conseguir una ganancia hedónica o de cualquier otro tipo (p.ej. una mejora de la reputación). El fenómeno de la comprensión de los estados mentales ajenos, y la motivación prosocial asociada a esta comprensión, es una capacidad de la naturaleza humana, si bien no se activa con la misma intensidad en todas las situaciones, ni tampoco se manifiesta psicológica y neurológicamente de forma análoga en todas las personas. La capacidad de resultar motivado para comportamientos de ayuda a otros individuos (emparentados o no) tiene una causación filogenética y ontogenética, independiente de la capacidad de cálculo anticipatorio³⁷².

La psicología y la neurociencia cognitivas contemporáneas consideran la simpatía una capacidad evolucionada que produce motivaciones genuinamente altruistas

³⁷² Mark Davis (1983, y 1994) ha realizado estudios basándose en su Índice de Reactividad Interpersonal (IRI), que muestran que las diferentes expresiones psicológicas de la comprensión y motivación de estados mentales de otros individuos obedece a una distribución estadística poblacional, en la que interfieren factores explicativos de tipo genético y ambiental. Por otra parte, los estudios realizados con gemelos socializados en ambientes familiares y culturales diferenciados acreditan que existe una base genética significativa en la explicación del predominio de actitudes y comportamientos pro o anti-sociales (Kendler, *et. al.* 2007).

(Eisenberg 2000; Preston y De Waal 2002a). La simpatía es un programa cognitivo-afectivo-operativo que resulta activado por la presencia de estímulos neurológicamente competentes, y que no puede ser objeto de operaciones utilitarias dirigidas a la búsqueda de una gratificación psicológica. Se trata de procesos neurofisiológicos automáticos, casi siempre inconscientes, y no elecciones deliberadas (véanse Decety y Chaminade 2003b; y Lamm, *et. al.* 2007). El individuo que simpatiza lo hace porque en su cerebro tiene lugar un proceso neural de alta complejidad, y lo lleva a cabo sin pasar previamente por la activación de los procesos ejecutivos conscientes. Esto implica, como manifestaron tanto Smith como Darwin, que la simpatía es un instinto, y no una preferencia aprendida que condiciona la elección³⁷³. Asimismo, la simpatía funciona selectivamente enfocando la atención ante situaciones de incumbencia del yo, y que pueden ser motivadoras de una acción *eficaz* del sujeto simpatizante en ayuda del sujeto simpatizado. En efecto, si el sistema neuroendocrino de la simpatía se activara indiscriminadamente, el organismo sería vulnerable al engaño y la explotación de oportunistas y depredadores (Loch, *et. al.* 2006: 229). Además, la activación persistente del sistema de alerta, ocasionada por la percepción reiterada de necesidades de ayuda sin ningún filtro discriminatorio, produciría una permanente situación de estrés y sobrecarga emocional (*compassion fatigue*) nociva para la salud y la eficacia del individuo simpatizante. Para evitar estas disfunciones, el sistema de la simpatía no opera de manera continuada ni indiscriminada, sino que lo hace a través de un proceso de filtrado selectivo que contiene procesos cognitivos y emocionales (Vignemont 2006). Como indican Andreoni y Rao (2011: 519), la especie humana está adaptada para un intenso altruismo compasivo, lo cual requiere la capacidad paralela para «ponernos anteojeras y protegernos de las tendencias altruistas naturales».

Pasemos a explicar la naturaleza de los estímulos que sí pasan el filtro de dichos procesos selectivos.

6.4.1.- La categoría instintiva del “nosotros”

Se ha demostrado que la percepción de la identidad de personas conocidas, integrantes de la familia, amistades, y otras que se suponen próximas y afines, por las

³⁷³ Un análisis del instinto de la simpatía, o altruismo compasivo, en estos dos autores se encuentra en Wispé (1991: 34 y ss.)

que el sujeto experimenta emociones vinculatorias, produce la activación inconsciente de la respuesta galvánica de la piel³⁷⁴. Por el contrario, las personas aquejadas de la enfermedad denominada prosopagnosia tienen dificultades de diverso grado en el reconocimiento de la identidad facial. Otro tipo de patología relacionada con el reconocimiento de la identidad de las personas es el llamado síndrome de Capgras, un cuadro delirante de tipo paranoide que afecta a la percepción de la familiaridad. Los pacientes con esta enfermedad no *pre-sienten* dermogalvánicamente la identidad de sus allegados, y por ello no reconocen los integrantes de su red de vinculación fuerte: su presencia suele ser calificada por ellos mismos de desconocidos o impostores (Hirstein y Ramachandran 1997).

En esta línea de estudios experimentales, cuando se somete a personas a la visión de imágenes con rostros de amigos, celebridades de la vida social que resultan valoradas y agradables, e individuos conocidos pero sin vinculación afectiva, se registra en el primer caso una sobreactivación de las redes neurales de la empatía y de los circuitos de recompensa (amígdala – hipocampo, núcleo accumbens, y córtex prefrontal ventromediano), mientras que la activación desaparece o es mucho menos intensa en los dos casos restantes, siendo menos significativa en el caso de los famosos que en el de los conocidos no amigos (Güroglu, *et. al.* 2008).

Existe pues un diseño funcional evolucionado del cerebro específicamente adaptado para las interacciones sociales de alta confianza y cooperación. Los individuos que no expresan, o que lo hacen con baja intensidad, el alelo del gen moniamina oxidasa A (una variante asociada con el comportamiento violento), manifiestan ya durante la infancia y/o la adolescencia una disfunción patológica en el funcionamiento de estos circuitos neurales especializados en la gestión del comportamiento afiliativo y prosocial. Por otra parte, la operatividad de tales redes neurales tiene demostrados efectos a largo plazo, puesto que la neuroquímica asociada a su activación reiterada levanta el estado de ánimo y evita los procesos orgánicos causantes de la depresión clínica. Existe evidencia de que los individuos con un funcionamiento anormal del área

³⁷⁴ Saygin, *et. al.* (2012) han demostrado que el cerebro activa automáticamente procesos funcionales de reconocimiento de rostros y atribución de identidades -con resultados de motivación y comportamiento- en área neurológica del giro fusiforme.

neuroológica afiliativa son propensos a comportamientos antisociales y depresivos, dos fenómenos que tienen impacto negativo en la eficacia (Güroglu, *et. al.* 2008: 908)³⁷⁵.

Las investigaciones llevadas a cabo por Damasio (2004) en sus pacientes con lesiones neurológicas en los subsistemas modulares encargados de procesar la información y el comportamiento social, como la empatía y la comprensión de las normas sociales, muestran que tales dispositivos son esenciales para las relaciones interpersonales y la inclusión social. Un patrón de comportamiento no sujeto a normas y carente de rumbo emocional-social, y/o un comportamiento carente de empatía, ponen en peligro la supervivencia del individuo porque se dificulta la coherencia interna y externa de las acciones, así como el encaje armónico con los comportamientos de otros individuos en la interacción social. Así, los daños en circuitos de la amígdala o en lóbulos prefrontales producen dificultades de diverso nivel en las habilidades de relación interpersonal, asunción de compromisos y seguimiento de las normas sociales aunque, curiosamente, estos individuos pueden mantener sus facultades cognitivas a pleno rendimiento, incluso con cocientes de inteligencia lógica por encima de la media. En este sentido, el autismo, por ejemplo, es una patología incapacitante porque, aun mostrando la mayoría de los pacientes un cociente intelectual normal o alto, y un sistema perceptivo sano, sin embargo se manifiestan incapaces de realizar inferencias simples acerca de los estados mentales ajenos (Thornhill, *et. al.* 1997: 226; Tomasello, *et. al.* 2005: 686). Por el contrario, el síndrome de Williams es un trastorno congénito cuyos enfermos manifiestan retraso intelectual, pero sin embargo no encuentran dificultades en las relaciones sociales ni en la comprensión de los estados mentales ajenos.

Los procesos neurológicos de reconocimiento de la red de personas con que el individuo se identifica, y que considera un “nosotros” integrado, constituyen instintos previos a la deliberación, y ese reconocimiento pre-cognitivo es un estado emotivo que,

³⁷⁵ Otros beneficios asociados a la epigenética afiliativa tienen igualmente demostrados efectos benéficos en la eficacia individual. Así, entre otros, se ha detectado una alta correlación entre la amistad y los comportamientos amistosos con indicadores de integración social, supervivencia en la enfermedad del cáncer, bienestar en edades avanzadas, reducción de la secreción de cortisol y del estrés, y en general, una mayor esperanza de vida. Asimismo, se ha demostrado una asociación negativa entre la carencia de amigos y comportamientos amistosos con la superior mortalidad por cáncer, así como una superior prevalencia y recidivas de la depresión clínica y otras enfermedades mentales. Véase Güroglu, *et. al.* (2008: 903).

al reconocer familiaridad, predispone a acciones de tolerancia, cooperación y altruismo (Aknin, *et. al.* 2011). Los dispositivos innatos de lectura de otras mentes e identificación simpatética son la base epigenética de los comportamientos de ayuda y solidaridad, que son modalidades de altruismo compasivo, y también de castigo y represalia, tanto en individuos emparentados como en desconocidos. La evolución de los procesos de causación de estos procesos motivadores del comportamiento ha requerido, no obstante, de la evolución paralela de una psicología adaptada para el reconocimiento de identidades afines:

Los comportamientos cooperativos habrían resultado en ventajosas alianzas para aquellos capaces de señalar esta capacidad, y la eficacia mejorada o el éxito material resultante habría intervenido en la proliferación de los comportamientos cooperativos que son en sí mismos una señal. (Bowles y Gintis 2003: 437).

El extremo opuesto al autismo y a los trastornos antisociales, que podríamos denominar “promiscuidad empática”, resulta igualmente patológico e inadaptativo (Hoffman 2000: 214-215). En efecto, existen sesgos epigenéticos en la capacidad de la especie humana de empatizar con otros individuos y realizar actos de altruismo. Tales sesgos guardan relación también con la señalización de identidades y la eficacia inclusiva. Entre los más destacados figuran el favoritismo por los más conocidos, los miembros del propio grupo primario (familiar o cultural), los amigos, individuos fenotípicamente semejantes, y los que están presentes en una relación presencial de “aquí y ahora” (*ibid.*, pp. 206 y ss.)

Richerson y Boyd (2001) han denominado “instinto tribal” (*tribal instinct*) a esta propensión conductual humana de tipo afiliativo y altruista en la que se encuadrarían, entre otras, las conductas de liderazgo prosocial, sumisión jerárquica, solidaridad con los afines, compromiso y conformismo con las instituciones del grupo³⁷⁶. Una versión de la teoría del instinto tribal es la que Dennis W. Organ denomina “comportamiento de ciudadanía organizacional” (*organizational citizenship behavior*, OCB), o “síndrome del buen soldado”:

³⁷⁶ Adviértase que aunque existen argumentos teóricos, y experimentos de sociología de la organización y psicología social que avalan la validez de la hipótesis del instinto tribal, sin embargo, todavía no contamos con experimentación neurocientífica concluyente que verifique la hipótesis, y explique cómo operan los sistemas modulares del cerebro relacionados con la experiencia intersubjetiva y la motivación altruista en las relaciones “intra-tribales”. Véase Singer y Lamm (2009: 92-93).

... el comportamiento individual que es discrecional, no directa o explícitamente reconocido por el sistema formal de recompensas, que a nivel agregado contribuye a promover el funcionamiento efectivo de una organización. (Organ 1988: 4).

En sus experimentos con los procesos cognitivos implicados en el uso y atribución de significado de las categorías y símbolos demarcadores del grupo, y los comportamientos que los miembros del grupo consideran “propios” de la pertenencia al mismo, Tajfel, *et. al.* (1971) han demostrado que en los procesos de conformación espontánea de órdenes grupales intervienen motivaciones discriminatorias de los que no son miembros del grupo, así como una valoración sesgada pro-grupal, o favoritismo inclusivo. La psicología humana está adaptada para buscar justificaciones coherentes de las tendencias segregadoras y discriminatorias, sobre la base de argumentos racionalizadores de una “hipocresía moral” aceptada por los afines (Norton, *et. al.* 2004). Los comportamientos de pertenencia se refuerzan con la identidad grupal, que es una fuerza motivacional de inclusión (y a la par de segregación) sobre la base de marcadores grupales. El estímulo motivacionalmente competente consiste en creer

... que el grupo es una entidad única, importante y atractiva. La identidad de grupo “alza una frontera” alrededor que define claramente la afiliación y facilita sentimientos de inclusión y apego (...) es una creencia colectiva que favorece la sensación, entre los integrantes del grupo, de que sus objetivos y futuros están positivamente vinculados. (Druskat y Wolff 2005: 205).

La creencia en la *competencia* del grupo activa la selección grupal, dado que «proporciona a los integrantes el sentido de que pueden ser más efectivos conjunta que individualmente» (Druskat y Wolff 2005: 206). De ello se desprende que la cultura que consigue el éxito identificador de sus depositarios es en sí misma una capacidad idiosincrásica del grupo con valor adaptativo, un patrimonio común que motiva, *cæteris paribus*, al individuo a cooperar: (a) sin necesidad de incentivos explícitos, (b) sin necesidad de supervisión externa, y (c) aun admitiendo costes personales (efectos externos) no internalizados (Tyler y Blader 2001: 222-223). Obviamente, el “éxito cultural” implica la viabilidad del modelo de selección en varios niveles. La cultura, como capital social, es un bien comunal que pertenece al grupo, y cuyas especificidades carecen de valor fuera de esta organización. El individuo identificado con un grupo porta los rasgos culturales de ese grupo, y tales rasgos definen las fronteras de lo que le

resulta incumbente o no, de las emociones inclusivas y excluyentes, y por lo tanto de la operativa para una racionalidad grupal (*we-rationality*).

Dentro de este contexto, Gintis, *et. al.* (2003: 166) señalan que sólo mediante la naturaleza cultural humana es posible que el altruismo no nepótico característico de la especie haya podido evolucionar mediante procesos coevolutivos, en los que la selección de grupo no supone abnegación en favor de la colectividad. La clave está en las instituciones culturales y en la plasticidad fenotípica, que permiten al individuo conseguir resultados medidos en eficacia individual superiores a los que podría obtener en un modelo de selección sólo individual, o dentro de arreglos culturales alternativos; y ello sin necesidad de conflictos inter-grupales frecuentes, tamaño limitado de los grupos, y tasas de migración exiguas, que son los principales requisitos de un modelo teóricamente verosímil de selección grupal pero, como se sabe, escasamente descriptivo del grupalismo empírico en la especie *Homo sapiens*.

6.5.- Psicología evolucionada para la intersubjetividad

Ya se ha explicado que la especie humana dispone de una adaptación destacada única en el reino animal por sus propias características, consistente en su desarrollada capacidad de acceder a las experiencias mentales ajenas y simpatizar con los “propietarios” de las mismas. Este rasgo epigenético se despliega completamente durante la infancia, y ya en el transcurso de la adolescencia los individuos «llegan a ser especialistas en la lectura de las intenciones ajenas» (Blaffer 2009: 7). Tanto la capacidad para la intersubjetividad, como la del lenguaje, que son características y prominentes la especie, constituyen dos rasgos dentro de un paquete adaptativo más amplio que hace posible su naturaleza hiper-social. Estos rasgos permiten la comunicación apta para una sociabilidad cultural y una cooperación altamente enriquecidas, así como para detectar desviaciones del comportamiento que puedan poner en peligro el modelo de supervivencia cooperativa humana. No obstante, la activación de los procesos cognitivos y emocionales que activan las motivaciones proclives a los comportamientos prosociales está relacionada con la percepción de rasgos de humanidad, y dentro de éstos, la semejanza interviene aumentando la intensidad de tales procesos motivacionales:

Los individuos eligen pertenecer o participar en grupos sociales con características que creen que les aportan eficacia; así, eligiendo una imagen social que es la más aproximada, o que es la más coherente con la propia imagen personal de sí mismos. La coherencia o incoherencia -un tipo de disonancia cognitiva- entre la imagen personal y la imagen social, es lo que motiva las acciones y elecciones del individuo. (Davis 2006: 381).

En efecto, la semejanza de rasgos genéticos y/o culturales está demostrada como factor explicativo de la modulación de la intersubjetividad y la activación de comportamientos prosociales³⁷⁷. Rilling, *et. al.* (2004b: 1.698 y ss.) muestran que el comportamiento de cooperación secuencial en juegos de dilemas sociales de tirada única, realizados en condiciones de anonimato, se produce paralelamente a la activación de redes neurológicas relacionadas con la atribución al co-jugador de estados mentales y atributos psicológicos propios de la condición humana. Esto explica que los participantes cooperen un 32,8% más de las ocasiones con otros participantes humanos que cuando saben que están jugando con un artefacto tecnológico. Las áreas neurales de la empatía-identificación, e inferencia por analogía de los estados mentales ajenos, se activan intensamente siempre que los jugadores cooperan con contrincantes humanos, y en los casos en que se coopera con artilugios informáticos ocurre asimismo, aunque con mucha menor intensidad. Nada de esto ocurre, sin embargo, en los casos en que los jugadores eligen no cooperar. Asimismo, la reciprocidad negativa es ejercida como respuesta cuando el jugador contrincante no coopera, aunque la aplicación de reciprocidad negativa es mucho más frecuente en interacciones con un co-jugador humano que con el programa informático. Esto indica que el cerebro activa comportamientos de cooperación o represalia como si estuviera efectuando una premeditada selección inclusiva de cooperadores, siempre que se trate del cerebro de un cooperador del que surgen las decisiones³⁷⁸. Del mismo modo, el que se coopere con más frecuencia y más intensamente con jugadores humanos que con un artilugio

³⁷⁷ Entre esos rasgos compartidos expresados en el fenotipo Hoffman (2000: 294) cita como demostrados las preferencias, actitudes, intereses, objetivos y preocupaciones. No obstante, la psicología social reconoce una propensión al conformismo social que oculta discordancias, y que se manifiesta, por ejemplo, en que «la gente tiende a asumir que sus amigos son como ellos, cuando de hecho existen áreas de desacuerdo que simplemente no se discuten» (McPherson, *et. al.* 2001: 435).

³⁷⁸ La conducta cooperativa está motivada en una respuesta neurológica favorable al comportamiento o actitud (cooperativa o no) del otro, y no a un cálculo exclusivo de potenciales ganancias individuales. Esta respuesta emocional está alienada con la lógica de la eficacia inclusiva del rasgo cooperación. En consecuencia, la cooperación indiscriminada es inverosímil, puesto que la cooperación es una variable dependiente de la identidad (cooperativa o no) del co-jugador. Véase Brosnan (2008) para un análisis en primates no humanos, y Fehr y Fischbacher (2003) para el caso humano.

tecnológico también denota que el cerebro selecciona inclusivamente los aparentes rasgos de humanidad³⁷⁹.

Las adaptaciones filogenéticas relacionadas con la lectura de otras mentes, al igual que la transmisión del conocimiento a través de la imitación, son capacidades que pueden influir positivamente en la eficacia inclusiva, en el sentido de que unas y otras inducen al individuo a seleccionar aquellos otros individuos dotados de tendencia al comportamiento cooperativo, y cooperar con ellos de manera predecible y coordinada (Gallese y Goldman 1998: 495-496; McCabe, *et. al.* 2000). El sistema de neuronas espejo sitúa a los individuos que interaccionan un plano cognoscitivo comprensible de estados neurológicos semejantes, de modo que cuanto mayor semejanza exista entre las estructuras neurales, filogenéticas y ontogenéticas, más capacidad tendrán de “leerse” recíprocamente las intenciones y estados mentales. En la literatura sobre la cognición a través de neuronas espejo se denomina a este fenómeno *reflejo especular* (Iacoboni 2009). Téngase en cuenta de que en un hipotético caso de reflejos especulares perfectos (como jugar con el reflejo de uno mismo), los individuos dispondrían de toda la información sobre las intenciones de sus contrincantes. La estrategia dominante en estos casos, aun en juegos como el dilema del prisionero y jugadores canónicamente racionales, es cooperación. De este modo, en el caso ideal de existir información perfecta acerca de las intenciones ajenas la estrategia dominante pasa a ser cooperación incondicional³⁸⁰. Dicho en otras palabras: «Donde no hay incentivo para la defección unilateral, los agentes racionales pensarán y actuarán como un equipo» (Schmid 2007: 178). Asimismo, cuando los jugadores de un juego de dilema social, como el dilema del prisionero, toman decisiones con una intencionalidad común (*collective intentionality*), el dilema desaparece, dado que la función objetivo a optimizar no es el pago individual,

³⁷⁹ Los experimentos con técnicas de neuroimagen de Saxe y Kanwhiser (2003), Kilner, *et. al.* (2003) y Stanley, *et. al.* (2007) demuestran que las áreas neurológicas especializadas en la comprensión y la empatía involucradas en los comportamientos altruistas no se activan, o lo hacen de forma insignificante, cuando el objeto percibido no expresa rasgos o conductas propias de la naturaleza humana. Existe un sesgo instintivo pro-específico en las capacidades mentales para la cognición social y la motivación altruista.

³⁸⁰ Véase Howard (1988) al respecto. Como explica Binmore (2009: 119-121) los jugadores racionales elegirán sistemáticamente cooperación en el dilema del prisionero de tirada única cuando ambos tienen una única e idéntica función de utilidad, y además lo saben. Este es un caso extremo de cooperación incondicional que Binmore llama jugada de los gemelos idénticos, algo así como jugar con una réplica exacta de uno mismo. Lógicamente, en términos biológicos esto equivaldría a maximizar la eficacia inclusiva, dado que los jugadores están maximizando una misma función de utilidad, y la percepción psicológica de sus pagos individuales se ve transformada por la naturaleza de la interacción.

sino la combinación de pagos conjuntos del grupo o equipo (*ibid.*) La existencia de una función de utilidad grupal transforma el problema de una optimización estratégica individual en una optimización paretiana grupal. Así, el juego deja de ser un dilema social para transformarse en un juego de coordinación, pues el individuo no se comporta como un ente autónomo, sino como miembro de un grupo que tiene intenciones y persigue objetivos grupales³⁸¹.

Raimo Tuomela (2000: cap. 2) establece como factores integrantes de una actitud cooperativa (de equipo, o *we-attitude*) los siguientes: (1) el individuo tiene (o comparte) dicha actitud, (2) cree que otros también la tienen, y (3) cree que hay una mutua creencia compartida en que otros (que son los miembros del equipo) tienen tal actitud. Téngase en cuenta que la identificación con una actitud pro-grupal supone admitir la hipótesis de una racionalidad trans-personal (que puede ser de un grupo reducido, y llegar al nivel de especie), y puede surgir incluso de una manera espontánea entre personas desconocidas, sin previa organización formal, como cuando se aplaude o abuchea en un evento público multitudinario. Hay indicios experimentales de que existen áreas neurales especializadas en el procesamiento de la cognición social sincronizada, cuya operatividad produce la convergencia en la atención para el logro de una meta común (*on-line joint attention*) (véase Redcay, *et. al.* 2010).

Lógicamente, estas capacidades tienen valor de supervivencia, debido a que permiten extender la cooperación recíproca y el altruismo más allá de la red familiar, mejorando así la eficacia inclusiva del rasgo cooperación. La clave que explica esta dinámica expansiva de la cooperación, que son las capacidades para la lectura de otras mentes y la comprensión intersubjetiva, también se beneficia de este proceso de mejora de la eficacia inclusiva, dado que es el rasgo principal que permite la expresión de los rasgos cooperativos y altruistas con desconocidos. Los individuos capacitados para leer mentes ajenas, y sentirse además afectados por los estados emotivos de los demás, podrán desarrollar acciones de eficacia inclusiva a través de la cooperación y del

³⁸¹ A este respecto señala Sugden (2000: 192) que: «Un equipo existe cuando sus integrantes se consideran a sí mismos miembros del equipo». Obviamente, esto se produce cuando se da una identificación de los roles con unas preferencias, objetivos, intenciones, creencias, etc., conjuntas, como en el caso de la identificación del individuo con las organizaciones formales y los movimientos sociales ya existentes, organizados en torno a determinados ideales, símbolos y objetivos (Bacharach 1999: 132).

altruismo exogámicos. Esta es la base de las ventajas asociadas a la especialización funcional y el comercio, así como de la colaboración social en pos de propósitos comunes y acción colectiva, conductas que caracterizan las instituciones de la sociedad extensa y de la propia civilización (Richerson y Boyd 2008).

6.5.1.- La perspectiva plural

Max Weber (1969[1922]: I, pp. 33-34) asoció la comunidad empírica a los sentimientos morales de identificación e inclusión social de sus integrantes. Según su punto de vista, lo característico de las *comunidades* es la cohesión motivada en afectos vinculatorios y compromisos con tradiciones, mientras que las *sociedades* resultan cohesionadas por la transacción de intereses individuales o la unión de esos intereses en estructuras de poder. En la comunidad la acción individual está grupalmente referenciada, pues el individuo se considera sentimentalmente parte de un agregado social de orden superior. Por el contrario, en la sociedad el individuo se identifica exclusivamente consigo mismo. Tuomela (2000) recupera esta idea y considera la naturaleza de las relaciones (cooperativas o no) en función de la actitud y mentalidad con que las afronta el sujeto: cuando éste persigue únicamente la optimalidad individual no está cooperando. Sí coopera cuando su objetivo es la optimalidad desde la perspectiva grupal. Esto supone una resolución armónica de los intereses particulares dentro de la multiplicidad de equilibrios posibles. Un reparto de 99 – 1 en el juego del dictador es óptimo en el sentido de Pareto, sin embargo no es óptimo desde la perspectiva grupal porque el nexo asociativo es inestable.

Dentro de esta línea analítica, Ferdinand Tönnies planteó en *Comunidad y asociación* una teoría de la identidad basada en la herencia dual, en la que las comunidades replican endógenamente hábitos de cooperación intracomunitaria a través de procesos sociales educativos y directivos:

... aparte de los instintos y fuerzas heredados, la influencia de una comunidad, como voluntad educadora y directiva, constituye el factor de determinación más importante en cuanto que atañe a la condición y formación de cada hábito y disposición individual. Sobremanera importante es el espíritu de familia, pero también lo es todo espíritu que se le asemeje y posea los mismos efectos. (Tönnies 1979[1887]: 45).

Según Tönnies las comunidades disponen de una forma de capital social que él denomina “consenso”, consistente en una cooperación tácita fundamentada en motivaciones pro-grupales que tienden a armonizar los posibles conflictos de intereses y dar estabilidad al grupo. Este consenso es «una fuerza y propensión social que mantiene unidos a los seres humanos como miembros de una totalidad» (Tönnies 1979[1887]: 45). La motivación del consenso es una capacidad psicológica innata dependiente del factor semejanza, que incluye las herencias genética y cultural:

El cimiento verdadero de la unidad, y, consecuentemente, la posibilidad de la comunidad, estriba, en primer término, en la estrecha relación consanguínea y la mezcla de sangre; en segundo término, la proximidad física y, por último, la proximidad intelectual. (Tönnies 1979[1887]: 47).

En esta perspectiva, para Tuomela (2007: 3) estos procesos de convergencia, que desembocan en la acción cooperativa, se traducen en lo que él denomina *we-perspective* en la toma de decisiones. El éxito de los mismos se ve influido por variables tales como el *ethos* del grupo, compromisos recíprocos interiorizados, emociones inclusivas (*we-feeling*), y acciones basadas en tales motivaciones. La diferencia entre el individualismo conductual y la perspectiva grupal estriba pues en que en la intencionalidad compartida el individuo identificado con un agregado social mayor se considera a sí mismo actuando no como individuo autónomo, sino en calidad de miembro de un grupo.

La psicología social moderna admite que la mayoría de las conductas prosociales fuera de los experimentos de laboratorio se llevan a cabo en ámbitos relacionales de tipo comunitario, dentro de patrones de acción habituales, planificados y estereotipados (esto es, de acuerdo a roles sociales), que hacen que las conductas de ayuda no sean conductas aisladas, ni realizadas prioritariamente con desconocidos sino que, *cæteris paribus*, y por condicionantes de tipo cognitivo y emotivo, están incrustadas en relaciones sociales personalizadas, con expectativa de reciprocidad incierta y a largo plazo (Akin, *et. al.* 2011; Amato 1990). De acuerdo a estos supuestos, la cooperación tiene fundamentalmente una explicación relacional e idiosincrásica. Por ello las características individuales (como los rasgos de personalidad, o los valores o ideología interiorizados por educación), y la circunstancias objetivas de la situación (como la urgencia de una necesidad presente), son variables explicativas, pero que predicen en

menor grado los comportamientos empíricos de ayuda que la adscripción a grupos sociales motivada por la identificación y la pertenencia, las relaciones a largo plazo, y la interiorización de las normas de funcionamiento interno de los grupos a los que se pertenece. El comportamiento prosocial humano viene explicado, tanto en su ejecución como en su intensidad, fundamentalmente, por variables de tipo relacional y vinculatorio (Amato 1990). ¿En qué consiste la causalidad próxima de esta disposición conductual? Se contemplan dos variables explicativas que están conectadas por el factor semejanza: (1) el vínculo de apego neuroquímico, y (2) la identificación por actividad del sistema de neuronas espejo. Se explican seguidamente.

6.5.2.- Neuroendocrinología de la afiliación y el vínculo

El vínculo de apego a personas y objetos se aprecia en la especie humana como una capacidad congénita de reconocimiento de, y preferencia por, lo considerado familiar y parte de sí mismo (Hood y Bloom 2008). La vinculación tiene un primer estadio en la proximidad relacional y reiteración de las relaciones (Zajonc 1968). De esto se derivan historias compartidas e identidades grupales, o bien esos grupos ya conformados activan en el observador su identificación por simpatía hacia estereotipos y categorías culturales que el individuo considera definitorias de sí mismo, y que observa en otros individuos semejantes a él (Tajfel 1972, y 1978). Dadas las circunstancias propicias, tales categorías pueden consistir en rasgos periféricos de la identidad personal que habitualmente pasan desapercibidos, mientras que en otras ocasiones forman parte del núcleo definitorio del autoconcepto³⁸². Con los desconocidos el vínculo primario puede surgir de la identificación con los rasgos de humanidad u otro rasgo periférico o circunstancial, y en cualquier caso a través de la empatía y las condiciones que definen el marco específico (*frame*) en que se produce la interacción (Fiske 2009).

La representación del dolor o la necesidad de ayuda expresada por otro, que se produce simultáneamente en los cerebros de quien sufre en primera persona y de quien

³⁸² Entre los rasgos que se ha demostrado pueden activar procesos identificatorios, De Waal (1997: 40) cita entre otros la edad y el nivel socioeconómico, preferencias políticas o religiosas, los antecedentes étnicos, el cociente intelectual, nivel cultural, atractivo físico y estatura. Tuomela (2007: 5) enumera algunos aspectos intervinientes en la identificación con un *ethos* grupal, tales como compartir metas y valores grupales, creencias y normas internas de comportamiento grupal.

observa (e incluso siente) empáticamente el sufrimiento o la necesidad (lo experimenta de forma vicaria), es el proceso neural que pone en alerta la disposición altruista. Ésta es modulada autónomamente en función de variables tales como la vinculación previa con la identidad del individuo observado, evaluación cognitiva de su situación y circunstancias personales, intensidad de la necesidad de ayuda o sufrimiento percibido, relación existente entre ambos sujetos, y experiencia anterior del observador en el sufrimiento, pudiendo resultar en una acción de ayuda realizada o inhibida³⁸³. Este sistema de la empatía es un proceso neurológico epigenético que posibilita la realización de actos de altruismo incluso con individuos desconocidos, y que caracteriza a la gran mayoría de individuos de la especie *Homo sapiens*.

Los afectos vinculatorios, y la amalgama de sentimientos que conducen a la identificación con las vivencias de otros yoes están causados por la acción autónoma del sistema nervioso central, cuando se activan las redes neurales de la empatía, y el proceso cognitivo y afectivo de identificación se ve reforzado por la liberación de la neurohormona oxitocina (en la mujer), y la vasopresina (en el varón)³⁸⁴. Se ha comprobado experimentalmente que los comportamientos prosociales de diversas especies animales guardan relación directa con la presencia de estas neurohormonas en el cerebro, y que las patologías del tipo de la fobia social o agresividad intraespecífica están asociadas al mal funcionamiento de los neuroreceptores de estos péptidos, así como a la deficitaria presencia de ellos en el cerebro y en el torrente sanguíneo (Austin 2006: 120-123). La endoproducción de oxitocina y vasopresina se explica por la filogénesis del comportamiento amoroso-vinculatorio asociado a la sexualidad y al cuidado de la progenie (Carmichael, *et. al.* 1994; López Moratalla y Sueiro Villafranca 2008). La liberación de oxitocina y su recepción por las células especializadas de la red ventromedial induce además la liberación de dopamina (un péptido opioide endógeno del sistema neurológico de recompensa) en el cerebro de determinadas especies de mamíferos, entre las que se encuentra la especie humana. En estas circunstancias, la

³⁸³ Véanse Hein y Singer (2008: 156), y Frith y Singer (2008: 3.877) para una alusión a trabajos relevantes acerca de las variables que intervienen en la modulación de la empatía.

³⁸⁴ La presencia de estas neurohormonas en el cerebro humano motiva comportamientos prosociales, dado que los receptores de las mismas se encuentran localizados en áreas asociadas a la experiencia emocional de acercamiento y tolerancia social (la amígdala, el hipotálamo y el cíngulo anterior) (Verbalis 1999). Asimismo, se ha constatado que el sistema se realimenta en la interacción social a través de relaciones de proximidad, conocimiento mutuo, contacto físico, seguridad relacional y señalización interpersonal de actitudes de confianza y no agresión (Zak 2005).

liberación de dopamina contribuye a generar sensaciones agradables asociadas a la percepción, relación y ayuda a parientes (Bartels y Zeki 2004; Liu y Wang 2003)³⁸⁵. Su evolución ha estado por tanto asociada a las ventajas de la eficacia inclusiva (como p.ej. su influencia en la reducción de la promiscuidad sexual, y el compromiso con la pareja en la crianza de la progenie), y por ello juega un papel fundamental en el comportamiento altruista de sacrificio en favor de la familia³⁸⁶. Curiosamente, el mismo patrón de comportamiento altruista que favorece la eficacia inclusiva puramente genética, resulta igualmente operativo y estimulado por el sistema de la oxitocina y la empatía fuera de los lazos de consanguinidad. Así, el vínculo afectivo con los individuos más vulnerables y necesitados de ayuda para sobrevivir, como es el caso de los niños pequeños, se ve reforzado no sólo con sus madres y padres biológicos, sino también, aunque en menor grado, en los padres y madres adoptivos, y en general en todas aquellas personas que han tenido contacto íntimo o frecuente con el niño (Champagne y James 2005).

Se constata además que el sistema de liberación de estos péptidos afiliativos, y otros con efectos similares como la arginina, no se activa solamente en las relaciones familiares, sino que ocurre también en las relaciones exoparentales: (a) relaciones grupales del tipo de proximidad relacional o comunitaria, (b) caso de relaciones esporádicas con desconocidos que señalizan una identidad étnico-cultural afín, (c) e incluso en casos de interacciones con desconocidos que no señalizan afinidad alguna, pero que manifiestan necesidad objetiva de ayuda (Ebstein, *et. al.* 2009; Carter 1998; Barraza y Zak 2009). Los fundamentos de buena parte del comportamiento social humano, especialmente en las estructuras sociales afiliativas exogámicas, tienen su

³⁸⁵ Sobre el funcionamiento del sistema neuroendocrino de recompensa en la liberación de péptidos como la dopamina, serotonina y endorfinas véase Damasio (2005: 64 y 113). La liberación de estos péptidos cumple una función adaptativa, relacionada con la reducción del dolor y del sufrimiento psíquico en circunstancias favorables a la supervivencia y reproducción (p.ej. el parto materno y ayuda en la superación de enfermedades sufridas por parientes). La bioquímica del bienestar psicológico interviene igualmente en camuflar el estrés psíquico causado por la aversión a la pérdida de recursos de supervivencia en favor de otros (véase Moll, *et. al.* 2006).

³⁸⁶ Adviértase que la maternidad, y curiosamente también la paternidad, producen transformaciones permanentes en la sensibilidad de la amígdala de los progenitores, que provocan una mayor sensibilización de su sistema de alerta y disposición a la ayuda, provocadas por el estímulo del llanto de su bebé. Esta respuesta es mucho mayor que la producida por el estímulo de la risa de éste. La sensibilidad reforzada permanece funcionalmente activa aún con los estímulos procedentes de bebés que no son sus hijos/as. Sin embargo, el estímulo de la risa del bebé activa con más facilidad la atención del individuo receptor que no tiene relación parental con él y no ha sido anteriormente padre o madre. Véase el estudio de Seifritz, *et. al.* (2003).

origen en los sentimientos de apego vinculatorio e identificación empática causados por la oxitocina y la vasopresina. Los estados mentales prosociales asociados a estas experiencias, inducidos por los dispositivos cognitivos y emocionales de la identificación y el apego afectivo, están reforzados, pues, por la activación de un sistema neuroendocrino exadaptado, diseñado originariamente para la eficacia inclusiva familiar (Pedersen 2004).

Las neurohormonas afiliativas no sólo desempeñan un papel en la modulación de los comportamientos altruistas y prosociales en general, en la tolerancia a extraños y la reducción del estrés y la agresividad, sino que además interfieren en la manera en que el cerebro procesa la información social. En efecto, en la gama de experimentos realizados en la última década con inhalación de oxitocina Kosfeld, *et. al.* (2005) hallaron que los individuos a los que se administra oxitocina sintética desvían su conducta social respecto al grupo de control que recibe un placebo. Los primeros se mostraron empíricamente mucho más confiados que los segundos en dilemas sociales como el juego de la confianza. Sin embargo, ese cambio de comportamiento no se ve acompañado de un aumento de su confianza psicológica en los demás individuos, que se manifiesta similar a la expresada por los participantes tratados con placebo. En comparación con éstos, los individuos que inhalan oxitocina no experimentan una alteración significativa de su comportamiento ante estímulos objetivos de altruismo, si bien manifiestan una mayor disposición a estabilizar relaciones de confianza y generosidad.

El experimento llevado a cabo por Zak, *et. al.* (2007) con los juegos del ultimátum y del dictador muestra que los individuos a los que se administra oxitocina se comportan con mayor desprendimiento que los no tratados, si bien, como en el experimento de Kosfeld, *et. al.* (2005), la oxitocina no altera los sistemas neurales encargados de procesar el comportamiento altruista, y por lo tanto no altera la disposición de los participantes que la han inhalado hacia un altruismo extraordinario. Estos resultados informan de que la oxitocina interviene alterando las preferencias (*cæteris paribus* aumenta la confianza y la generosidad), pero no las creencias (respecto al valor de la confianza o la generosidad, la necesidad de ayuda del otro, la fiabilidad, o

la identidad de los demás)³⁸⁷. Obviamente, esto podría hacer a los individuos con mayor disposición a activar el sistema de las neurohormonas afiliativas más vulnerables al riesgo de explotación por los que no tienen esa capacidad congénita; y ello aun sabiendo que pueden ser objeto de trampas y engaños, dado que las creencias sobre los demás no se ven alteradas por esta hormona. Esta consecuencia no es un asunto de importancia menor, como vamos a explicar.

En el experimento de Zak, *et. al.* (2007) los participantes tratados con oxitocina finalizaron su participación en el juego con una acumulación neta de pagos monetarios inferiores en un 5,2% a la media de los que obtiene el grupo tratado con placebo, lo que demuestra que la activación del sistema de la oxitocina tiene un coste medido en eficacia biológica. Esto plantea un desafío en términos de análisis evolutivo de este dispositivo neuroendocrino. En efecto, el sistema neurológico de regulación de estos péptidos sólo puede haber evolucionado si existe algún sistema paralelo que dificulte o compense la pérdida de eficacia, como pueden ser procesos neurológicos inducidos de cribado selectivo de potenciales vinculaciones. Esto puede consistir en que el sistema se active de manera selectiva en relaciones con individuos cargados de replicadores (biológicos y culturales) semejantes, como es el caso de la preferencia por el etnocentrismo (De Dreu, *et. al.* 2011). Y dentro de este contorno ya de sí restringido, otro cribado selectivo puede ser la activación preferentemente en relaciones especulares, o sea, con individuos igualmente endoproductores de oxitocina, y por lo tanto destinatarios de una vinculación recíproca.

Estas hipótesis pueden confirmarse por el hecho constatado de que el gen que sintetiza la proteína que hace posible la recepción de la oxitocina por el sistema

³⁸⁷ Aunque los estudios con oxitocina administrada todavía no han producido resultados concluyentes, y diversas hipótesis están en fase de contrastación, sin embargo, se ha constatado por Singer, *et. al.* (2008) que la oxitocina no influye en la activación del circuito de la empatía que explica un mayor altruismo (ínsula anterior – córtex cingulado anterior), sino en la reducción de la estimulación de la amígdala, que explica la mayor confianza y desprendimiento. En efecto, la oxitocina refuerza la confianza y la generosidad no porque el sujeto empatice con la situación dolorosa o desagradable de otro, sino porque al funcionar como “quitamiedos” (reduce la excitación de la amígdala, encargada de anticipar estados ansiosos de temor y preocupación), ello facilita al individuo desprenderse de recursos para entregárselos a otro individuo necesitado, y también sentir el estrés empático con menor intensidad, lo cual facilita la confianza en la aproximación a personas y situaciones problemáticas. Curiosamente, como indican Singer, *et. al.* (2008) las conductas de pretendido egoísmo racional en los experimentos de neuroeconomía, pueden estar más relacionadas con disposiciones muy emotivas (la sobre-excitación de la amígdala, y la experiencia neurótica de miedo y desconfianza asociada), y no tanto con actitudes carentes de emotividad del tipo cálculo racional desapasionado.

nervioso está involucrado igualmente en la mayor destreza en el reconocimiento facial de identidades (Skuse, *et. al.* 2014). Adicionalmente, existen pruebas de la destreza de los individuos para la selección de relaciones basadas en la reciprocidad de emociones afiliativas, es decir, la preferencia por emparejamientos con individuos inclinados a producir simultánea y recíprocamente neurohormonas de vinculación. Zak (2005) ha demostrado que los rasgos responsables de la vinculación y confianza a través de la producción de oxitocina, y los que causan experiencias empáticas, se replican simultáneamente (entre otras razones porque están correlacionados) a través de los comportamientos de cooperación con semejantes que portan estos mismos rasgos neurológicos. De este modo, los individuos que reciben de personas desconocidas actitudes y acciones de confianza y cooperación, muestran mayoritariamente una activación neural inducida del sistema de la oxitocina con una intensidad proporcional a la intensidad de la actitud y grado de confianza y cooperación recibidas. Como consecuencia, el 98% de los participantes en el experimento respondieron con un nivel de endoproducción de oxitocina altamente correlacionado con el nivel que supuestamente sería la causa de las actitudes y comportamientos que ocasionan esta respuesta de gratitud. De este modo, se puede admitir la hipótesis del círculo virtuoso de la oxitocina, o la realimentación de la cooperación a través del programa neurológico instintivo de la oxitocina. Subráyese, no obstante, que esto es válido para el 98% de la población participante en el estudio³⁸⁸.

Esto nos sitúa en un modelo de eficacia inclusiva convencional, en el que el individuo provisto de capacidad genética para el altruismo afiliativo está además capacitado para discriminar a los que no portan ese rasgo recíprocamente. Si eso es así, la evolución del altruismo a través de la neurobiología de la oxitocina resulta verosímil. El factor semejanza resulta relevante para explicar la evolución de los comportamientos vinculatorios. Consecuentemente, la preferencia por la semejanza fenotípica habría

³⁸⁸ Resulta conveniente pues ampliar el estudio y profundizar en sus variables explicativas. En cualquier caso, puede advertirse que, curiosamente, la proporción de individuos aquejados de trastornos antisociales de tipo psicopático es algo más del 2% (Hare 2006). Estos individuos no disponen de la capacidad congénita de la empatía, si bien su sistema de oxitocina no necesariamente es patológico. Consecuentemente, aunque estos individuos sí podrían responder cooperando sobre la base de su sistema de oxitocina, al no ser capaces de empatizar, no pueden *sentir* lo que motiva a los simpatizantes a cooperar, y por ello en su cerebro se cortocircuita la realimentación de la confianza social.

evolucionado para que la vinculación, y las conductas prosociales asociadas entre cooperadores, tengan viabilidad evolutiva.

6.6.- Marcadores fenotípicos, semejanza y *rapport*

El fenómeno del *rapport* se asocia conceptualmente al acoplamiento, conexión, resonancia mental entre individuos, o reflejo neural, que permite comprender sin intervención del raciocinio lo que el otro está experimentando subjetivamente. Esta conexión puede llegar a ser tan intensa que influya en constantes vitales como el ritmo cardíaco y respiratorio, sudoración, electricidad corporal, ondas cerebrales, liberación de hormonas y neurotransmisores, así como liberación de endorfinas en la sangre y tejidos del sistema nervioso (Goleman 2006: cap. 1). El *rapport* activa circuitos neurológicos de atención sincronizada que posibilitan la acción conjunta a través de la cooperación; esto se da en mayor medida en interacciones que se producen simultáneamente (cara a cara) y en tiempo real (Redcay, *et. al.* 2010). Tal capacidad de sincronización y acoplamiento fisiológico consiste en una adaptación evolucionada para la sociabilidad compleja, que incluye el desciframiento de estados mentales e intenciones ajenas, la inferencia y atribución de intenciones grupales, y la computación de efectos externos derivados de las acciones propias y ajenas (Tooby y Cosmides 1996: 129-130). Este patrón de sociabilidad compleja se halla adaptado para la interacción tanto con individuos conocidos y estables, como con desconocidos (Tomasello, *et. al.* 2005). Así, salvo casos patológicos (como la psicopatía y el autismo) o de individuos especialmente entrenados para engañar simulando estados mentales e intenciones,

... la sinceridad es la respuesta por defecto del cerebro (...) nuestro sistema nervioso transmite todos los estados de ánimo a la musculatura facial, evidenciando de inmediato nuestros sentimientos. Este despliegue emocional es automático e inconsciente, razón por la cual su represión exige un esfuerzo consciente y deliberado. Es por ello que tratar de distorsionar lo que sentimos y de ocultar el miedo o el enfado exigen un esfuerzo que rara vez consigue completamente su objetivo. (Goleman 2006: 19).

Del mismo modo que la sinceridad está codificada como respuesta en la comunicación no verbal, el cerebro humano dispone de adaptaciones específicas para descifrar la información verbal y no verbal, evaluar su coherencia y contextualización, y proponer hipótesis sobre estados subjetivos ajenos, evaluaciones y estimaciones de intenciones y

conductas³⁸⁹. Esta adaptación interviene de forma decisiva en la conformación de juicios sobre la fiabilidad de los otros, y en la ejecución, o no, de conductas de cooperación y altruismo (Pinker 2004: 520).

El dispositivo neural específico encargado de llevar a cabo el *rapport* es el sistema de neuronas espejo, emplazado en su mayor parte en el área pre-frontal del cerebro. Existen grupos de neuronas especializadas en decodificar patrones motores, y descifrar el significado de los movimientos simples y acciones complejas observadas en otros. La comprensión se lleva a cabo porque estas neuronas reproducen a través de un *reflejo especular* la activación visceromotora y emocional de los comportamientos percibidos en otros individuos. La operación es completamente automática:

... una comprensión que aparece desprovista de toda mediación reflexiva, conceptual y/o lingüística, al basarse únicamente en ese *vocabulario de actos* y en ese *conocimiento motor* de los que depende la capacidad de actuar. (Rizzolatti y Sinigaglia 2006: 125; cursivas en el original).

La epistemología de la vida mental ajena tiene como base el etograma, así como la gama de reflejos y patrones motores innatos, pero también destrezas aprendidas, comportamientos estereotipados, creencias y normas culturales interiorizadas, experiencias vividas, conceptos y teorías asimiladas, etc., codificado en el cerebro del individuo observador a través de la interacción con el ambiente social y cultural. Así pues, las neuronas espejo pueden evocar una comprensión de la vida mental ajena tanto más precisa en la medida en que se disponga de un patrimonio de codificaciones motoras, emocionales, lingüísticas y de significado cultural semejantes. Lógicamente, la pertenencia a una misma especie, y dentro de esto, a una misma cultura; y dentro de la cultura, haber tenido experiencias ambientales y haber desarrollado aprendizajes de destrezas y conocimientos técnico-teóricos similares, conecta a los individuos mediante un *rapport* mejor acoplado y más exacto, un *espacio de acción compartido* que posibilita una empatización inmediata, más precisa e intensa, en la que se *capta* en

³⁸⁹ Place, *et. al.* (2009) proponen que este desarrollo cognitivo está relacionado con la adaptación previa de la capacidad de descifrar las intenciones del cortejo sexual, así como la formación de hipótesis sobre la respuesta positiva o negativa al cortejo. Ese circuito epigenético puede haberse exadaptado para descifrar valoraciones e intenciones ajenas en toda la interacción social. De ahí que los animales sociales sean más diestros en la lectura de intenciones ajenas de los individuos de su especie, sobre todo las especies que se reproducen sexualmente y a través de sistemas de cortejo ritualizado.

segunda persona el significado de la vida mental ajena (Rizzolatti y Sinigaglia 2006: 134-137). Así, la experiencia indica que se capta sin ningún esfuerzo intelectual extraordinario el mensaje emitido por un comentarista radiofónico que se expresa en la lengua materna del oyente. Sin embargo, el cerebro no capta información comprensible alguna cuando se escucha eso mismo en una lengua desconocida o, siendo legos en etología animal, cuando se escucha el canto de un pájaro o de una ballena. Lo mismo se puede decir acerca de los rituales de comportamiento social percibidos en otras especies de animales, o simplemente en individuos humanos que pertenecen a una cultura desconocida. Cuanto mayor es el patrimonio de codificaciones neurológicas compartido entre el observador y los individuos observados, ya sea éste cultural o genético, del individuo y de la especie, mayor es *el rapport* y consecuentemente el espacio de comprensión y acción compartidos de manera espontánea.

6.6.1.- Homofilia: cooperación entre cerebros semejantes

El sistema de neuronas espejo genera experiencias empáticas capaces de motivar la acción a través de la percepción o imaginación de expresiones físicas y movimientos del otro. Con esta información se produce el reflejo especular, en el que: (a) se activan las neuronas espejo que provocan la simulación interna de la información recibida, (b) la ínsula conecta estas áreas de neuronas espejo con los centros cerebrales del sistema límbico, y (c) el sistema límbico activa los procesos emocionales que son el eco sincronizado de los que probablemente se están produciendo en el individuo observado o imaginado (Iacoboni 2009: 163)³⁹⁰. El lector puede producir una experiencia de comprensión empática observando las ilustraciones de la figura 6.3., en las que se muestran expresiones no verbales de estados emocionales en individuos humanos y no humanos. El cerebro descifra espontáneamente y sin esfuerzo cognitivo el significado de emociones, disposiciones y actitudes básicas que verosíblemente vivencia el individuo percibido, porque se hallan codificadas en el observador y en el objeto que este percibe.

³⁹⁰ Las disfunciones patológicas en el sistema de neuronas espejo cortocircuitan la activación del sistema límbico y los procesos emocionales en las personas con autismo. Véase al respecto el trabajo de Oberman y Ramachandran (2007).

Figura 6.3.- La experiencia transversal de comprensión empática



Fig. 6.3.1.- Al realizar la mueca de amenaza, el mandril baja las comisuras de los labios de modo que se aprecie en toda su longitud el colmillo superior. Los humanos hacen lo mismo, a pesar de que sus colmillos bastante más cortos. Arriba, a la izquierda, mandril; en el centro, actor de kabuki representando una actitud colérica. Derecha, muchaha furiosa. En Eibl-Eibesfeldt (1994: 19).



Fig. 6.3.2.- A la izquierda, muchaha samburu coqueteando: contacto visual, descenso de los párpados y desviación de la mirada. A la derecha, niña alemana de tres años que esconde el rostro detrás de las manos con juguetona coquetería. En Eibl-Eibesfeldt (1994: 49 y 50).



Fig. 6.3.3.- Indígena de la Gran Nicobar saludando con la mano abierta en alto. Derecha: niña napolitana saludando con la mano abierta en alto, al mismo tiempo que abraza protectora a su hermanito. En Eibl-Eibesfeldt (1994: 186 y 187).



Fig. 6.3.4.- Expresión mímica característica de tristeza y preocupación.
En Frank (2008a).

Que el sistema de neuronas espejo active un reflejo especular en los circuitos neurales del observador, y que las estructuras neurales co-activadas sean las responsables de codificar la misma gama de repertorios motores, cognitivos y emocionales compartidos entre individuos observador y observado no implica, sin embargo, que las redes neurales co-activadas sean idénticas, ni que se acceda en segunda persona una experiencia idéntica a la que vivencia el otro, lo cual, dicho sea de paso, es imposible. En efecto, la complejidad del cerebro humano es de tal nivel que incluso gemelos homocigóticos educados en un mismo ambiente no disponen de cerebros (y por tanto, de posibilidades de experiencia) idénticos³⁹¹. No obstante, una superior semejanza fisiológica entre los cerebros, ya esté explicada por la proximidad genética o por la neuroplasticidad desarrollada en la ontogénesis, permite una superior calidad en las operaciones de sincronización, simulación e imitación que posibilita el sistema de neuronas espejo. Esto mismo comprende una mayor calidad de la empatía, y por tanto de la identificación entre individuo observador e individuo observado, ya que la experiencia, y no sólo las programaciones innatas, predispone y condiciona la actividad de las neuronas espejo (Rizzolatti y Sinigaglia 2006: 136-137). La afinidad filo y ontogenética son variables explicativas del nivel de *rapport* que activa los comportamientos de cooperación y altruismo. Así lo explica Iacoboni (2009: 221),

... la actividad del sistema de neuronas espejo es un índice de nuestro sentido de afinidad con otras personas (...) estas neuronas nos ayudan a entender las acciones de los otros simulando en el cerebro las mismísimas acciones con la activación de nuestros propios planes motores. Al hacerlo, también nos ayudan a sentir lo que sienten otras personas (...) estas células parecen crear cierto tipo de “intimidad” entre el yo y el otro, y tiene sentido

³⁹¹ Como explica Iacoboni (2009: 148): «No existen dos cerebros con el mismo tamaño, forma o estructura interna, ni la activación de la misma área cerebral en distintas personas es nunca idéntica».

postular que la actividad del sistema de neuronas espejo también puede estar vinculado con el sentido de pertenencia o de afinidad respecto de un grupo social específico cuyos miembros, sentimos, son más similares a nosotros que otras personas.

Lógicamente, la activación de los circuitos neurales especulares permite compartir en segunda persona una experiencia ajena, y este fenómeno psicológico genera sentimientos de pertenencia e incumbencia recíprocos entre las personas y sus circunstancias personales y sociales. Parafraseando al neurocientífico Vittorio Gallese, lo que se vivencia es «como si el otro se transformara en otro yo», una descripción que apunta a lo expresado por el filósofo Maurice Merleau-Ponty cuando dice: «es como si la intención del otro habitara mi cuerpo, y la mía, el del otro»³⁹². Pero dado que la activación de los circuitos especializados en la empatía es un programa selectivo (a fin de evitar el estrés empático), el cerebro discrimina automáticamente en la producción del reflejo especular de las neuronas espejo. Algunas variables que intervienen en este proceso de selección de señales que activan la identificación son: (a) la identidad reconocida en el otro, (b) el tipo de relación que se tiene con él/ella, (c) las circunstancias del observador en el momento presente, (d) el grado de conocimiento de la persona observada y de su reputación, (e) la historia compartida de interacciones, (f) y la evaluación cognitiva sobre la justicia y moralidad de las intenciones del individuo observado (Rizzolatti y Sinigaglia 2006: 182-183). En definitiva, intervienen condicionantes puramente económicos (el coste/beneficio real del altruismo que pueda tener lugar), así como identitarios (¿qué rasgos de identidad biológica o cultural se comparten con el posible beneficiario?), de reciprocidad (¿existe posibilidad o no de que la ayuda sea devuelta en alguna ocasión futura?), y de contexto (externalidades e interdependencias del tipo intereses y/o amenazas comunes).

El concepto sociológico de homofilia consiste a grandes rasgos en el principio según el cual el contacto entre individuos semejantes tiene lugar con una frecuencia e intensidad superiores a la que se produce entre individuos diferentes (McPherson, *et. al.* 2001: 416). La psicología experimental ha demostrado que la atracción interpersonal resulta afectada por la similitud percibida (Huston y Levinger 1978). Tal y como menciona Lewis (2005: 708), para un yo la semejanza es

³⁹² Ambas citas se encuentran en Iacoboni (2009: 81).

... una importante propiedad del mundo en el que vive, en el que llega a ser parte del conocimiento cultural, la transmisión de ideas, y la causa de atracciones y aversiones.

Por otra parte, se ha constatado que la genética determina de manera significativa la selección de socios en la interacción social, o sea, la propensión a buscar semejanzas fenotípicas (condicionadas por la genética) en los individuos candidatos a formar coaliciones y grupos estables de cooperación (Kendler, *et. al.* 2007; Güroglu, *et. al.* 2008).

La interpretación biológica y económica de este fenómeno social se basa en que la relación con afines ha de aportar alguna ganancia comparativa extraordinaria, a corto o a largo plazo, al individuo y/o a su *pool* de replicadores (biológicos y culturales). En términos biológicos, la búsqueda de semejantes por parte de cooperadores (*positive assortment*) es una condición favorable a la evolución de la reciprocidad y del altruismo (consanguíneo o no) (Bowles y Gintis 2008: 10). Además permite explicar procesos de selección en varios niveles (Hamilton 1975; Simon 1983: 60-61) y selección cultural-grupal (Richerson y Boyd 2005).

En términos económicos, las preferencias sociales como la valoración de la justicia, el interés por el bienestar ajeno, el igualitarismo, las expresiones del compromiso con una identidad, etc., requieren marcos institucionales que reduzcan los costes de transacción de localizar cooperadores, en orden a reducir costes de búsqueda y evitar la explotación y el oportunismo en la acción colectiva (Tooby, *et. al.* 2006). Hasta instancias insospechadas, la existencia de cooperación requiere una capacidad evolucionada de comunicación que posibilite detectar afines y señalar la afinidad, pero este proceso es costoso aun con un marco institucional propicio. Una vez superado este problema -aún no solucionado o no solucionado del todo- de la búsqueda y detección de semejantes, existen argumentos suficientes para defender la hipótesis de que la semejanza facilita tanto el altruismo por la vía de la eficacia inclusiva directa y a través de la *simpatía*, como la cooperación simbiótica entre afines, una forma de eficacia inclusiva indirecta a través de la *empatía cognitiva* y la *mentalización*, que facilita la coordinación de la cooperación recíproca, así como la detección y evitación de oportunistas y depredadores (Hoffman, *et. al.* 1998; Sally 2000). En ambos casos la semejanza facilita la coordinación y calidad de la cooperación, el establecimiento de

metas comunes y la ejecución de planes de acción que requieren una mentalidad de equipo (Tomasello, *et. al.* 2005; Sugden 1993, y 2000).

A nivel cognitivo, los individuos que comparten *pool* de replicadores biológicos y culturales tienen la ventaja de contar con las bases neurológicas necesarias para una identificación empática y simpatética de mayor calidad y más motivadora, de la cual surgen externalidades positivas propias de la simbiosis de semejanza:

... esperamos que las personas se asocien con otras similares a ellas dada la mayor facilidad de la comunicación, los gustos culturales compartidos, y otras peculiaridades que facilitan la coordinación de la actividad y la comunicación. (McPherson, *et. al.* 2001: 435).

... la mayor eficacia de los filtros de identidad (contribuye) a equilibrar el mayor tamaño del grupo (y sus ventajas) con una reducida heterogeneidad cultural de sus integrantes; esto mejora la efectividad de la supervisión recíproca y la construcción de una reputación grupal, que contribuyen a incrementar los altos niveles de cooperación intra-grupal... (Estos beneficios compensan en) los grupos más homogéneos su abstención de aprovechar las economías de escala, ganancias del intercambio, y los posibles beneficios derivados de la difusión del conocimiento asociada a la presencia de una mayor y más diversa composición de miembros integrantes del grupo. (Bowles y Gintis 2003: 438; paréntesis añadidos).

La predisposición a cooperar depende en gran medida de la creencia acerca de la identidad de los individuos con quienes se interacciona, y esta disposición universal, denominada “localismo” (*parochialism*), puede haber coevolucionado a través de procesos de selección en varios niveles (Bowles y Gintis 2003: 437-438). La delimitación del grupo a través de marcadores fenotípicos de identidad favorece la cohesión y la solidaridad internas, evita la entrada de oportunistas, y señala la disposición competitiva frente a grupos rivales. Según Cosmides y Tooby (2004), las motivaciones conducentes a la característica identificación humana con marcadores fenotípicos, y predominantemente culturales como la etnia o la religión, no tiene un instinto propio que la active, sino que se trata de un subproducto de otro dispositivo neuropsicológico que ellos denominan “rastreo de alianza potencial” (*alliance tracking*). La especie humana cuenta con esta adaptación psicológica, consistente en la capacidad de detectar rasgos potencialmente aptos para la cooperación dentro de alianzas viables y productivas:

Las pistas arbitrarias (tales como el color de la piel) cobran -o pierden- significación en la medida que tienen validez predictiva para la membresía de una coalición. (Cosmides y Tooby 2004: 120).

Habida cuenta que la especie humana manifiesta una elevada diversidad cultural surgida de su particularmente alta plasticidad fenotípica, la similitud de rasgos (genéticos y culturales) fenotípicamente expresados es un factor primordial de clasificación en función de categorías sociales e identidades grupales. La comunicación y la señalización veraces pasan a ser capacidades psicológicas evolucionadas, tanto por el lado de la emisión como por el de la recepción de información relacionada con los rasgos de identidad. Ello se explica porque los agrupamientos de afines que comparten identidad aportan ventajas selectivas a quienes disponen de las destrezas necesarias para llevarlos a cabo. En esta perspectiva, cobra sentido el argumento filosófico de Hannah Arendt: «Al comunicar los sentimientos, los placeres y goces desinteresados, se expresan las preferencias y se eligen las compañías» (Arendt 2003[1982]: 136).

6.6.2.- Cognición cultural y altruismo en alianzas y amistades

Según explican Tomasello, *et. al.* (2005: 687-688) el éxito de la hominización vino caracterizado por la moderación o supresión adaptativa de algunos de los instintos de intolerancia y agresividad que caracterizaron al ancestro común de los homínidos pre-humanos, hace 6 millones de años. Estas adaptaciones neurológicas permitieron una actitud relacional más amistosa, basada en compartir emociones, y en motivaciones para compartir experiencias, conocimientos, normas de comportamiento social, creencias, técnicas y planes de acción socialmente coordinados. Es decir, los pasos hacia la hominización consistieron en avances hacia una naturaleza social y grupal que requería una mente social adaptada para absorber y propagar replicadores culturales. Esto impulsó la progresiva evolución hacia el cerebro *sapiens*, una transformación asociada a procesos coevolutivos dentro de nichos culturales y trinquetes evolutivos:

... la hominización biológica fue necesaria para la elaboración de la cultura, pero la emergencia de la cultura fue necesaria para la continuación de la hominización hasta *neandertal* y *sapiens*. (Morin 2004: 33).

El rasgo distintivo del género *Homo* debió ser, en distintos grados, cierta disposición a la competencia exogrupal, compatible con «habilidades y motivaciones para colaborar unos con otros en actividades que requieren metas comunes y atención e intenciones conjuntas» (Tomasello, *et. al.* 2005: 687). Estas nuevas capacidades

aportaron ventajas adaptativas decisivas para la evolución humana, dado que permitieron acometer coordinadamente proyectos conjuntos de magnitud y complejidad únicas en la naturaleza. Tales disposiciones fueron posibles por la evolución de una capacidad mental presente en distintos grados en el reino animal: la racionalidad de grupo (*we-rationality*).

La capacidad de crear y construir conjuntamente instituciones culturales, aptas para la supervivencia social-grupal, requiere una previa adaptación para concebir objetivos y planes conjuntos, y ejecutarlos sin elevadas (e informacionalmente costosas) exigencias de comunicación y coordinación (Tomasello, *et. al.* 2005: 682). La coevolución de cultura y genética fue estableciendo progresivamente trinquetes evolutivos que refinaron esta competencia humana para la construcción de una cultura (material y simbólica) de carácter plural y público. Conforme las culturas fueron evolucionando en demes no hibridados, el factor semejanza fenotípica fue haciéndose crecientemente relevante para la evolución de la cooperación, hasta el punto de constituirse la etnia en una causa de exclusión frente a los *outsiders*³⁹³. La ventaja de disponer de una modalidad cultural de cognición se asocia a la evolución de motivaciones proclives al favoritismo por lo semejante, y dado que la semejanza decisiva en la construcción social de un nicho de supervivencia es la cultura, por ello la especie humana se halla capacitada para desarrollar conductas de cooperación y altruismo más allá de los miembros pertenecientes al linaje familiar.

Efectivamente, en la especie humana la percepción de semejanza activa motivaciones de inclusión y exclusión. Tales motivaciones están asociadas a la eficacia relativa de las posibles alianzas que puedan existir o llegar a ser entre los individuos (Levine y Kurzban 2006)³⁹⁴. La eficacia de las alianzas viene determinada, *cæteris paribus*, por el grado de semejanza fenotípica entre los individuos que interactúan:

³⁹³ Simplemente el hecho de utilizar una lengua, rituales de relación, o técnicas diferentes, ya son factores de exclusión para la cognición cultural, y por lo tanto para una cooperación espontánea y enriquecida.

³⁹⁴ Téngase en cuenta que la semejanza puede representar paralelismos y antecedentes homólogos. Es conocida la sagaz inteligencia social de los primates, que hace a estos animales especialmente aptos para seleccionar a los socios y colaboradores más leales, discriminar a los sospechosos de intenciones oportunistas, y valorar la calidad más que la cantidad en las relaciones de cooperación (Silk 2007; Noë y Hammerstein 1994). La cooperación y el altruismo de los primates sociales, no obstante, y a diferencia de la especie humana, son comportamientos de carácter fundamentalmente nepótico (Brosnan 2008; Blaffer 2009).

cooperación con los que permiten inferir una productiva alianza; indiferencia pasiva, o competencia por los recursos, con aquellos otros con quienes es inverosímil una alianza mutuamente beneficiosa. Los marcadores (individuales y grupales) que invitan a considerar una alianza viable son discrecionales, y sujetos al dinamismo de circunstancias ambientales, sociales, culturales e históricas (Caporael, *et. al.* 1989: 693; Loch, *et. al.* 2006: 225). Lo relevante aquí es que esos marcadores de semejanza resulten una *proxy* fiable de disposiciones fenotípicas que permitan una cooperación relativamente provechosa respecto a otras posibilidades de acción individual y colectiva³⁹⁵. No obstante, el marcador de semejanza primordial para establecer alianzas que refuercen la eficacia a través de la cooperación es la propia capacidad de cooperar, y la aproximación exploratoria requiere la emisión previa de señales fidedignas, así como la inferencia acertada de intenciones por los individuos intervinientes:

Ante un desconocido nos hallamos en una relación ambivalente, dudando entre simpatía y miedo, sin saber si va a mostrarse amigo o enemigo. Para pacificar la relación e ir hacia la amistad, intercambiamos gestos de cortesía con él. Pero estamos dispuestos, en caso de hostilidad, a huir, defendernos o atacar. (Morin 2004: 81).

En este sentido, Cosmides y Tooby (2004) presentan las coaliciones como grupos familiares ficticios o cuasi-grupos familiares. La exadaptación del instinto altruista de eficacia inclusiva familiar parece evidente, e incluso se puede decir que en la especie humana esta exadaptación llega a permitir fraguar alianzas a veces más fuertes y comprometidas que las de la consanguinidad, dado que pueden aportar mayor eficacia biológica al individuo capacitado para llevarlas a cabo que cooperar exclusivamente en el seno del linaje familiar (Loch, *et. al.* 2006: 224-225)³⁹⁶.

³⁹⁵ De acuerdo a Kanazawa (2001: 1.148) los marcadores más fidedignos en las sociedades modernas, por caracterizar rasgos identitarios más difíciles de tergiversar, son: raza, etnia, cultura, lengua y religión. Esto es así por ser los más estables, y lo son porque se trata de marcadores, a diferencia de otros como la ocupación, ideología política, o preferencias de ocio, adquiridos fundamentalmente por transmisión procedente de los grupos en los que el individuo se socializa en su infancia. Estos marcadores son más estables que otros, por tratarse de *imprintings* fenotípicos, que también pueden tener una base hereditaria epigenética. Véase al respecto Cavalli-Sforza, *et. al.* (1982).

³⁹⁶ En el modelo que propone Whitmeyer (1997) resulta ventajoso para la eficacia genética del individuo ayudar a otros individuos con quienes pueden construirse alianzas eficientes (presentes o a futuro), como por ejemplo el matrimonio propio o matrimonios de descendientes genealógicos. Las alianzas matrimoniales conllevan implícitamente alianzas en términos materiales y culturales, de producción y de defensa frente a amenazas externas, y las sinergias en estos aspectos mejoran la eficacia genética de las alianzas sexuales.

Admitido esto, para que la cooperación fuera del linaje familiar pueda evolucionar es preciso que se den adaptaciones psicológicas para señalar y captar rasgos fenotípicos afines con suficiente calidad y precisión, incluso soportando un coste a corto plazo en pos de una eficaz señalización (Gneezy, *et. al.* 2012). La conducta de cooperación está asociada en la especie humana a la activación de los circuitos especializados en la lectura de otras mentes y comprensión de intenciones ajenas (la corteza media prefrontal, donde se emplazan abundantes redes modulares de neuronas espejo). En los experimentos de neuroimagen realizados por McCabe, *et. al.* (2001) se aprecia que estas redes neurales de la intersubjetividad se activan en las personas que cooperan, y no se activan en aquellas otras cuya conducta no es cooperar. Asimismo, dicha activación se produce en los cooperadores en mayor medida cuando cooperan con otras personas, y en bastante menor medida cuando lo hacen con un artefacto informático. Esto quiere decir que durante la interacción el cerebro del cooperador está procesando información social relevante para cerciorarse de las posibilidades de que surjan resultados mutuamente beneficiosos, una forma de asegurar la eficacia inclusiva del rasgo cooperación a través del altruismo pasivo:

... la cooperación requiere una zona de convergencia activa que concentre la atención en la ganancia recíproca capaz de inhibir suficientemente la gratificación asociada a la expectativa de una ganancia inmediata mayor. (McCabe, *et. al.* 2001: 11.834).

La confianza pues se basa en una alta probabilidad de que el rasgo epigenético de cooperación exista en el destinatario de la misma, y ésta ha de ser tomada en cuenta por razones de eficacia inclusiva, ya sea a través de la reciprocidad entre cooperadores, o directamente a través del altruismo llevado a cabo de uno a otro de ellos. Una población étnica y culturalmente homogénea, en la que abundan los cooperadores, es un ambiente rico en capital social debido a la comunidad de externalidades positivas (*neighbour effect*), y a que el rastreo de alianza potencial (*alliance tracking*) deriva en probables emparejamientos, y acciones colectivas, entre cooperadores con un bajo coste de búsqueda. Ello posibilita una cooperación abundante, que ya sea en su vertiente de reciprocidad, o de altruismo, contribuye a incrementar la eficacia de ese rasgo. Así lo explica De Waal:

... los investigadores pueden predecir estadísticamente, basándose en lo bien que encajan las dos personas en diversos aspectos, si una pareja que acaba de conocerse romperá o seguirá junta. Al igual que en los monos rhesus, es probable que estas reglas de emparejamiento estén relacionadas con las posibilidades de cooperar: cuantos más rasgos e intereses una persona comparta con otra, más fácil les será llevarse bien y más probabilidades tendrán de establecer una relación de concesiones mutuas. (De Waal 1997: 40).

Según este enfoque, a los individuos les interesa encontrar a sus afines porque con ellos la cooperación es más productiva en términos de eficacia. Esto se debe a que entre afines existen relaciones de reciprocidad (a corto y a largo plazo), sinergias y externalidades positivas, además de altruismo³⁹⁷. Por otra parte, una red social homogénea, precisamente porque entre los afines existen externalidades positivas y otras sinergias, constituye una expresión informal de seguro y garantía recíproca, apta para la eficacia en condiciones de incertidumbre.

Tal y como argumentan Tooby y Cosmides (1996), la vinculación amistosa entre individuos no emparentados genealógicamente permite plantear una teoría del altruismo no-nepótico, y que además tiene la particularidad de que no debilita la eficacia (genética) del individuo a largo plazo, sino que la acrecienta. El problema evolutivo subyacente al modelo es el que ellos denominan “paradoja del banquero”, y consiste en que las motivaciones exclusivamente centradas en la eficacia individual, como el egoísmo racional en teoría de juegos, no cuentan con apoyos teóricos favorables en una perspectiva psicológica evolutiva del fenómeno relacional de la amistad. La relación de amistad, según los autores, es la que se caracteriza por la ayuda incondicional cuando ningún otro individuo acudiría a cooperar por reciprocidad. Por eso la “paradoja” del banquero:

... justo cuando los individuos necesitan dinero más desesperadamente, es cuando tienen un riesgo de insolvencia muy elevado, y por lo tanto, cuentan con la menor probabilidad de ser seleccionados para recibir un préstamo. (Tooby y Cosmides 1996: 131).

La conducta racional de evitar la selección adversa en el negocio bancario hace inverosímil ayudar al que está en una situación verdaderamente apurada. Y

³⁹⁷ Subraya Tuomela (2000: 13) que un ambiente de externalidades positivas en el que esté restringido el ingreso de *free-riders* es en sí un ambiente cooperativo, propicio para el surgimiento y la sostenibilidad de una preferencia por el bien común, aun en el caso extremo de que los individuos sólo actúen motivados por sus objetivos y preferencias individuales (es decir, sin identidad social).

curiosamente, este tipo de situaciones, al límite, hubieron de ser muy frecuentes en el ambiente evolutivo donde se produjeron las adaptaciones psicológicas que definen la actual naturaleza humana:

Para los cazadores-recolectores, la enfermedad, las heridas, la mala suerte en la caza, o la incapacidad para resistir un ataque de los antagonistas sociales, debieron ser frecuentes reveses de la fortuna con un impacto selectivo de primer orden. (Tooby y Cosmides 1996: 132).

En consecuencia, la evolución de procesos neuropsicológicos que contribuyeran a forjar entre los individuos compromisos de cooperación y altruismo, creíbles e incondicionales, constituyó un avance socio-adaptativo que mejoró exponencialmente la eficacia de quienes se mostraron capaces de ganarse la simpatía de otros, y de mostrar sentimientos de simpatía hacia los que cuya mera existencia reporta una ventaja superior, en términos de eficacia, que la vida solitaria o una vida social exclusivamente motivada por el egoísmo:

La habilidad para atraerse la asistencia durante esos amenazadores reveses de la fortuna, donde la ausencia de apoyos puede resultar mortal, bien puede haber tenido consecuencias selectivas más significativas que la capacidad para cultivar relaciones de intercambio social que promueven incrementos marginales en los réditos durante el tiempo en que el individuo está sano y salvo, y bien alimentado. (Tooby y Cosmides 1996: 132).

El problema adaptativo asociado a la paradoja del banquero pudo ser solucionado con la adaptación de simbiosis a largo plazo reforzadas por vinculación afectiva, en las que se producen comportamientos de altruismo sin una necesaria reciprocidad a corto plazo, aunque, a diferencia del altruismo materno, se trata de «un altruismo sujeto a evaluación periódica» (Tooby y Cosmides 1996: 140). Lo que hoy llamamos amistad consiste en el apego interpersonal motivado por emociones representadas en una función asimétrica de valor (Kahneman, *et. al.* 1991), como la que enlaza al individuo con el conjunto de sus propiedades con “valor sentimental”. En las relaciones de este tipo los individuos cuentan con un repertorio de capacidades (una identidad) específico y difícilmente sustituible. Por ello surge un refuerzo consistente en el recíproco celo de vigilancia y protección, que neuroquímicamente viene explicado por neurohormonas y neurotransmisores de vinculación afectiva. La bioquímica del apego intersubjetivo conlleva la necesidad de una privilegiada exclusividad relacional, y sentimientos

recíprocos de confianza y lealtad incondicional. De este modo, el objeto de vinculación acaba siendo un valor intrínseco y exclusivo para su valedor³⁹⁸.

Esta teoría de la amistad permite explicar la característica propensión humana a una mejora de la propia condición consistente en cultivar una identidad colectiva integrada en grupos, pero igualmente personalista y exclusiva (Tooby y Cosmides 1996: 133-134). Los humanos se esfuerzan y compiten por diferenciarse de los demás, quieren ser únicos, diferentes, exclusivos e imprescindibles para al menos alguien:

Gran parte de la vida social consiste en un continuo movimiento para encontrar y ocupar nichos personalizados que sean extraordinariamente beneficiosos para alguien y difíciles de imitar, acompañado esto de un entramado de relaciones sociales que busca configuraciones donde las partes sean mutuamente supervaloradas. (Tooby y Cosmides 1996: 134).

Obviamente, la exclusividad idiosincrásica de estos nichos personalizados hace que sean limitados en la vida de las personas, y ello tiene dos implicaciones: (a) no cabe la defección con los amigos, porque la ganancia de la estrategia racional puede suponer la pérdida de un “recurso” insustituible; y (b) dadas las limitaciones cuantitativas en la cantidad de estos nichos relacionales, los individuos han de «afrontar el problema computacional de llenar esos huecos con individuos de los que se pueda esperar los mejores resultados a largo plazo» (Tooby y Cosmides 1996: 136)³⁹⁹. Consecuentemente, la vida relacional consiste, en este aspecto, en el despliegue de habilidades para seducir al otro, y para no ser seducido por impostores (“falsos amigos”), sino por los mejores amigos posibles⁴⁰⁰. ¿Y qué requisitos tienen estos nichos relacionales, estos amigos

³⁹⁸ Eberhard Weisman defiende que además del grupalismo básico de muchas especies sociales, existe en la naturaleza animal otro instinto adicional de grupalismo exclusivo y altamente vinculador, que sobrepasa la funcionalidad del grupalismo genérico. Aplicando su hipótesis al comportamiento de muchas especies de aves monógamas, se pregunta (Weismann 1994: 216): «¿Acaso existe junto a los “instintos sociales” conocidos otro que no conocemos bien todavía?». A lo que contesta aludiendo al comportamiento sexual monógamo de los ánsares comunes:

... junto a todos los demás “instintos sociales” tiene que haber, además, una especie de instinto de relación, una “predisposición para la *simpatía específica*”. ¿Es quizás este instinto de relación el principio y el fundamento sobre el que se basa, sustancialmente, el comportamiento social de especies animales monógamas? (Cursivas añadidas).

³⁹⁹ Téngase en cuenta que el modelo planteado no requiere un reconocimiento inmediato de los rasgos fenotípicos del amigo óptimo. Como indican los modelos de Pacheco, *et. al.* (2008) y Yamagishi, *et. al.* (2005), la aproximación poco arriesgada entre individuos desconocidos a través de incipientes relaciones de cooperación recíproca puede conducir, a largo plazo, tanto al refuerzo de la lealtad cooperativa a través de sentimientos de vinculación específica, como al descubrimiento de un capital relacional específico del tipo que plantean Tooby y Cosmides.

⁴⁰⁰ El capital relacional de un individuo está en función de la cantidad de buenos amigos, y de los beneficios en cascada derivados de las externalidades positivas asociadas a esas redes de alta densidad

verdaderos, que aportan -y no restan- eficacia a largo plazo aunque la relación presuponga potenciales compromisos de altruismo a corto plazo? La clave está en una amplia semejanza fenotípica, que permite aprovechar todo tipo de sinergias relacionales de carácter idiosincrásico, como el *rapport*, economías de escala, complementariedades, gustos y creencias compartidos, externalidades positivas; objetivos, proyectos y amenazas comunes, etc., que reportan un “valor de asociación” superior a otras fórmulas de reciprocidad motivadas por la conveniencia utilitaria (Tooby y Cosmides 1996: 136-138).

6.6.3.- Construcción cultural de rasgos de humanidad y universalización de la benevolencia, o sea, de la predisposición al altruismo general

En 1922 Max Weber expuso que

... muchos afectos reales (...) y las reacciones derivadas de ellos, podemos “revivirlos” afectivamente de modo tanto más evidente *cuanto más susceptibles seamos de esos mismos afectos*; y en todo caso, aunque excedan por su intensidad a nuestras posibilidades, podemos comprenderlos endopáticamente en su sentido, y evaluar intelectualmente sus efectos sobre la dirección y los medios de acción. (Weber 1969[1922]: 7; cursivas añadidas).

Y Adam Smith explicó un siglo y medio antes que Weber:

Cada facultad de un ser humano es la medida con la cual juzga la misma facultad en otro. Yo evalúo la vista de usted según mi propia vista, su oído por mi oído, su razón por mi razón, su resentimiento por mi resentimiento, su amor según mi amor. No tengo ni puedo tener otra forma de juzgarlos. (Smith 1997[1759]: I, i, 3. p. 66).

La susceptibilidad de una misma experiencia afectiva, según Weber, permite una mejor comprensión y evaluación de los otros yoes, o sea, una intersubjetividad más precisa y motivante para la acción prosocial. Smith, por su parte, argumenta que las experiencias nunca son idénticas, sino el resultado de la intervención de determinadas facultades

relacional, y las conexiones con otras redes que conectan con el núcleo de mejores amigos (*network externalities*). De esta manera,

... haciéndose uno mismo valioso para otro, de esta manera pasa a ser también valioso para aquellos que consideran valioso a aquél y se preocupan por su bienestar. Así, los que se benefician del bienestar de ese otro también se benefician del bienestar del primero, porque ese bienestar afecta al bienestar de aquél. Por ello, el uno causa un beneficio *primario* al otro; el uno también beneficia a aquellos que se benefician del bienestar del otro como efecto *secundario*. Aquéllos, por este motivo, están interesados en el bienestar del primero, y esto conduce a beneficios *terciarios* que revierten directamente a él. (Levine y Kurzban 2006: 181; cursivas en el original).

psíquico-orgánicas que tienen los individuos y que, siendo semejantes entre ellas, permiten un resultado igualmente semejante del cual se puede inferir el conocimiento de la subjetividad ajena. Ambas intuiciones se alinean con las investigaciones científicas contemporáneas.

En efecto, el modelo teórico de percepción-acción (*Perception-Action Model*, PAM) presentado por Preston y De Waal (2002b) explica la empatía como producto de un sistema neurológico de representación de experiencias ajenas que resulta activado, lo hace con mayor rapidez, y opera con más exactitud motivadora para la acción cuando el individuo observador tiene (y percibe) semejanzas fenotípicas con el individuo observado (*ibid.*, pp. 16-17)⁴⁰¹. El isomorfismo que determina procesos mentales y conductas en paralelo y sincronizadas es un factor relevante en la explicación de la intersubjetividad. Obviamente, las semejanzas fenotípicas son una *proxy* de semejanzas epigenéticas, dado que el grado con que los animales interaccionan sobre la base de la empatía (dentro y entre las especies) está relacionado con la semejanza en morfología y biomecánica de las especies. Una vez superado el escollo de las diferencias interespecíficas de origen genético, los factores fenotípicos que explican la activación y precisión especular de la experiencia empática isomórfica están asociados a factores (directamente percibidos, imaginados, o fiablemente señalizados) locales-relacionales-culturales, entre los que los autores resaltan: (a) la familiaridad (resultar conocido, reiteración de las interacciones pasadas), (b) similaridad, (c) experiencias pasadas compartidas, y (d) aprendizajes similares llevados a cabo⁴⁰². Estas variables influyen en la calidad de la representación de los estados mentales ajenos, y una mayor calidad determina

... más asociaciones, y de este modo, crea un patrón de actividad en el sujeto (observador) más complejo, elaborado y preciso; este patrón se codifica en relación tanto con la experiencia personal como la experiencia (relacional) con el objeto. (Preston y De Waal 2002b: 16; paréntesis añadidos).

⁴⁰¹ Mitchell, *et. al.* (2005) han llevado a cabo diversos experimentos con técnicas de neuroimagen que permiten validar preliminarmente ésta hipótesis principal del modelo de percepción-acción (PAM).

⁴⁰² Un trabajo neurocientífico posterior, que reitera la validez explicativa de estos grupos de factores en la capacidad motivadora de la empatía para el altruismo, es el de Singer y Lamm (2009). No obstante, los autores concluyen que es preciso profundizar en la investigación del papel desempeñado por estas variables, y otras aún no estudiadas suficientemente.

Según el PAM, la clave para que se produzca una identificación empática (*affective congruence*, o *rapport*) suficientemente motivadora para activar comportamientos de cooperación y altruismo, y mentalidad grupal, está en el grado de semejanza neuroanatómica, neurofisiológica y neuroquímica que existe entre los cerebros. Neuroanatomía, neurofisiología y neuroquímica tienen una causación genotípica, y también un desarrollo relacionado con la ontogénesis, es decir, con la experiencia, el aprendizaje y la influencia ambiental durante la vida de los individuos:

Para evitar el conflicto y las pérdidas de utilidad, la gente puede querer agruparse con aquellos con quienes comparte la misma identidad, o con quienes sus acciones tienen un mismo significado. Así, el proceso de agrupamiento en sí mismo -las “situaciones” en las que los agentes se encuentran a sí mismos- puede ser endógeno, conducido por valores e identidades. (Akerlof y Kranton 2000: 732).

En la especie humana la empatía no surge mecánicamente de la observación de un individuo cualquiera en cualquier circunstancia, sino que otras variables culturales y relacionales, además de las biológicas, operan para que en la empatización cognitiva se lleve a cabo una evaluación de las características y circunstancias del sujeto percibido (Lamm, *et. al.* 2007; Singer, *et. al.* 2006). Asimismo, la identidad atribuida a los individuos percibidos es un factor que condiciona la activación de los procesos cognitivos y emotivos asociados a los procesos empatizantes y, dentro de estos, el grado de simpatía afectiva que puede surgir de esa evaluación de las circunstancias (Singer, *et. al.* 2004a). Tanto en uno como en otro caso, las creencias respecto a la condición de semejanza o desemejanza del sujeto observado influyen decisivamente en la activación de los procesos mentales propios de la humanidad de lo humano, y esa información, que condiciona el sesgo con que se percibe al otro, es una construcción cultural, que alude a la identidad del observador en relación con la del sujeto observado. Dice Amartya Sen al respecto:

El sentido de identidad puede ser fuente no sólo de orgullo y alegría, sino también de fuerza y confianza. No resulta sorprendente que la idea de identidad reciba una admiración tan amplia y generalizada, desde la afirmación popular de amar al prójimo hasta las teorías del capital social y la autodefinición comunitaria. Y, sin embargo, la identidad también puede matar, y matar desenfrenadamente. Un sentido de pertenencia fuerte -y excluyente- a un grupo puede, en muchos casos, conllevar una percepción de distancia y divergencia respecto de otros grupos. (Sen 2007: 23).

Los conflictos intergrupales, que vienen acompañados de emociones primarias como el odio, el miedo y la ira, forman parte de los procesos de selección en varios niveles para los que está programada la naturaleza humana. Se trata de conflictos culturales alimentados por creencias, y en cuya explicación fallan las teorías utilitaristas del individualismo metodológico, tal como explica Knight:

Los hombres muestran tendencia a formar grupos cuya unidad interna se basa en buena parte en un sentimiento de oposición o competencia con respecto a otros grupos, más que en la eficiencia económica o el interés por la pura fraternidad (...) Lo importante es un sentimiento capaz de procurar la necesaria cohesión, que parece siempre obedecer a un motivo externo; todo sentimiento de unidad parece implicar otro de diferenciación. (Knight 1976[1934]: 329).

Las creencias influyen en la categorización social y en las emociones ligadas a la cognición de una categoría cultural, o sea, en las preferencias⁴⁰³. Una estrategia de exclusión y segregación bien conocida consiste en propagar una creencia de diferenciación (genética, étnica, espiritual) hasta arraigarla en expresiones emocionales de la población, generalmente a través de un estilo de comunicación muy emotivo (*hot thinking*), con mensajes lanzados de manera que se implanten acriticamente en el sistema límbico sin una previa evaluación cognitiva (Simon 1983: 8-10)⁴⁰⁴. Con un programa de este tipo, que persigue el objetivo anti-cooperativo de despreciar y destruir al adversario, puede llegar a desactivarse la propensión natural humana al altruismo pasivo. Un ejemplo es el siguiente tomado del libro de Chirstopher R. Browning sobre la *praxis* de la llamada Solución Final, ejecutada por el régimen nazi en la Polonia ocupada:

El 27 de junio, después de entrar en la ciudad de Bialystok, el comandante Weis ordenó a su batallón que peñara el barrio de los judíos en busca de varones, pero no especificó qué era lo que tenían que hacer con ello. Por lo visto, eso se dejaba a la iniciativa de los capitanes de compañía, que habían sido orientados en su manera de pensar en la reunión previa a la invasión. La acción comenzó con un progromo: golpes, humillaciones, barbas quemadas y

⁴⁰³ Las creencias como variable explicativa de las preferencias es un tema ampliamente tratado en por la psicología social y la teoría de la elección pública (véanse Batson, *et. al.* 1987; Miller 1999, y Marwell y Ames 1981), y más recientemente por la economía experimental (Ferraro, *et. al.* 2005; Frank, *et. al.* 1993a). Esta causación forma parte del canon lógico característico del pensamiento racional, no así la relación inversa (Binmore 2009: 245). La racionalidad ecológica de las creencias vendría de los procesos de selección cultural, y a largo plazo podrían insertarse en el etograma vía adaptaciones coevolutivas en el sistema límbico.

⁴⁰⁴ En psicología este proceso de desactivación de las emociones reguladoras del comportamiento social, a través de la desinhibición de los comportamientos de odio y agresión, se explica por el condicionamiento al estímulo.

disparos a discreción mientras los policías conducían a los judíos hacia el mercado o la sinagoga. Cuando varios líderes judíos aparecieron en el cuartel general de la División de Seguridad 221 del general Pflugbeil y se arrodillaron a sus pies suplicando la protección del ejército, un miembro del Batallón Policial 309 se bajo la bragueta y orinó encima de ellos al tiempo que el general se volvía de espaldas. (Browning 2002: 44).

Estos comportamientos incivilizados, de innecesaria humillación sin honor ni ganancia alguna para quien los realiza, ocurren si se consigue desconectar la identificación básica con los rasgos de humanidad a través de la inhibición, por creencias activadoras de emociones poderosamente motivadoras como el odio, la ira y el miedo, de los automatismos empáticos y de teoría de la mente que conectan las subjetividades individuales⁴⁰⁵. Ello se consigue, por ejemplo, si se propaga la creencia en que un determinado grupo no tiene naturaleza humana (deshumanización), o se trata de una aberración de la misma (infracomunización) (Fiske 2009). Tales actuaciones van orientadas a cosificar despectivamente al otro mediante su despojo de toda posibilidad de relación intersubjetiva⁴⁰⁶. Por seguir con la temática ya iniciada (aunque caben numerosos ejemplos similares), uno de los horrores políticos del s. XX, basado en tales premisas condicionantes de las preferencias, fue el programa genocida de odio racial y darwinismo bélico diseñado por el aparato de comunicación del régimen nazi. La llamada Solución Final fue una estrategia aplicada para conseguir la cosificación, primero, y la estigmatización y satanización, más tarde, de la etnia judía:

Los nazis atribuyeron a aquellos elegidos para su exterminio ser “diferentes” (culturalmente) y además desviados (genéticamente) del alemán “tipo”. Tal vez les resultara más fácil destruir a aquellos considerados diferentes, practicar la barbarie mostrando escasa o nula empatía. El resultado fue la ausencia de simpatía, y una tendencia creciente a hacer daño. (Schulkin 2002: 47; paréntesis añadidos).

Cuando la interiorización de creencias se lleva a cabo con éxito, se implanta un marcador emocional susceptible de expresarse vía preferencias y de ser replicado a

⁴⁰⁵ Las investigaciones llevadas a cabo por Bruneau, *et. al.* (2012) demuestran una inhibición de la respuesta en los dispositivos neurológicos de la empatía y de la mentalización ante el dolor o el sufrimiento ajenos cuando existen conflictos culturales activos, y no así, o en mucho menor grado, en el caso del dolor o sufrimiento de individuos pertenecientes a una cultura simplemente distinta. Los estímulos que activan la empatía, la mentalización, y las emociones conducentes a motivar comportamientos de cooperación y altruismo tienen que ver con la percepción de rasgos de humanidad, y la estimulación adicional vía rasgos culturales puede resultar activadora o inhibidora de comportamientos humanitarios en función de las creencias interiorizadas asociadas a dichos rasgos.

⁴⁰⁶ La creencia de que un interlocutor no es humano, o bien la despersonificación del mismo, produce resultados similares a los experimentos de interacción social llevados a cabo con personas autistas o aquejadas de trastornos antisociales por incapacidad empática. En estos casos la cooperación se reduce de forma muy notable, e incluso desaparece.

través de la transmisión cultural de valores, especialmente a través del aprendizaje de prejuicios dentro de ámbitos relacionales segregadores⁴⁰⁷. Así, podría decirse que el meme cultural se implanta (o no) en un cerebro social, como el humano, porque previamente hay algo que hace atractiva su entrada, y ese algo explica que el contenido no sea rechazado o simplemente olvidado (Bartra 2006: 102-103). La expresión de estas creencias grabadas con valencia emocional puede ser inhibida mediante los procesos de control ejecutivo de la mente. Sin embargo, los estudios llevados a cabo por Phelps, *et. al.* (2000), Phelps y Thomas (2003), y más recientemente por Norton y Sommers (2011), demuestran que dichos fenómenos de psicología social, tales como la extensión del racismo, permanecen latentes en forma de prejuicios inconscientes en la población durante largos períodos de tiempo, y ello a pesar de que se ha comprobado una importante disminución de las actitudes explícitas de racismo en la expresión de las preferencias⁴⁰⁸.

No obstante, por el mismo procedimiento de propagación de creencias y éxito en la interiorización de las mismas, en este caso a través de emociones de tolerancia e inclusión, y practicando procedimientos relacionales que permitan captar estímulos generales de humanidad con mayor impacto en las preferencias que los particulares marcadores de identidad cultural, la cultura puede utilizarse también para generar identificaciones e identidades supra-culturales, que activen eficazmente la expresión de los instintos propios de la humanidad de lo humano, esto es, una re-humanización de las relaciones humanas (Bruneau y Saxe 2012). Existen indicios neurocientíficos de que la neuroplasticidad humana permite educar tanto para una intersubjetividad integradora a través del amor y la compasión universal, como para una intersubjetividad represora basada en la moralidad de la culpa, el fanatismo y el castigo (véase Singer y Steinbeis 2009: 47-49). Los instintos inclusivos están programados para expandir el espectro de incumbencia en la atención prestada a estímulos externos, y repercuten en los dispositivos de empatía y teoría de la mente. En consecuencia, motivan

⁴⁰⁷ Véase Castro Nogueira, *et. al.* (2003: 158 y ss.) sobre el papel de los marcadores emocionales en la evolución cultural humana.

⁴⁰⁸ Esto quiere decir que la eficaz replicación cultural de un prejuicio deja huella emocional en la amígdala y en el hipocampo, y genera procesos mentales inconscientes que pasan inadvertidos (como las aversiones, fobias, sentimientos de desconfianza y rechazo visceral) a pesar de la inhibición de su expresión conductual. Téngase en cuenta que el hipocampo es un registro de memoria emocional, y que la amígdala procesa, entre otras, las reacciones de alerta asociadas a las emociones sociales de miedo y desconfianza, cuya expresión es fácilmente detectable mediante indicadores de ansiedad.

comportamientos benevolentes de distinto grado en función de las circunstancias subjetivas y objetivas que se den en la relación. Su expresión más reducida motiva para el altruismo pasivo que caracteriza lo que podemos llamar civilización y comportamiento civilizado⁴⁰⁹.

La funcionalidad adaptativa de la plasticidad fenotípica humana no deriva en, o no lo hace exclusivamente, el gregarismo con el grupo y la agresión dirigida a los que no pertenecen al mismo, sino que tal capacidad contribuye también al acercamiento cooperativo a través de la enculturación en sociedades reguladas por sistemas de valores y códigos morales universalistas. Entre ellos se incluyen la tolerancia y el respeto a la dignidad de los individuos que no forman parte de las redes primarias y de cooperación fuerte, ya se trate de grupos biológicos y/o culturales (Hoffman 2000: 293-298; Zak 2008b). Este es uno de los rasgos básicos definitorios de las culturas “civilizadas”, armonizadas por el altruismo pasivo que, como indican Gintis, *et. al.* (2003: 154), en la actualidad se definen por instituciones tales como la democracia y el gobierno representativo, el Estado de bienestar, y otras instituciones socio-políticas de cooperación organizada con desconocidos, pero que, con ser estos sistemas políticos excepcionales en la historia de la humanidad, sin embargo los valores y preferencias que los caracterizan son coherentes con los sentimientos morales que capacitan al individuo para la simpatía y la benevolencia, y refuerzan la coevolución de estas conductas (Richerson y Boyd 2008; Solomon 2008). El intercambio de mercado mismamente, una institución que Edgeworth presentó en su *Psicología matemática* como un estado de guerra, sin embargo, en visiones más realistas,

... aunque esté impulsado por el interés propio, aún depende en forma absoluta de una comunidad que comparta valores tales como la honestidad, la libertad, la iniciativa, la frugalidad y otras virtudes cuya autoridad no puede ser reducida por mucho tiempo a un gusto personal, como hace explícitamente la filosofía del valor positivista e individualista. (Daly y Cobb 1994: 50).

Los sistemas culturales universalistas e inclusivos se basan en la valoración positiva de categorías representativas de sentimientos morales, descriptivos de

⁴⁰⁹ Se puede definir la civilización como un sistema de relaciones entre individuos regulado por un conjunto de reglas de coordinación que permiten hacer del comportamiento cooperativo una estrategia evolutivamente estable. Colman y Browning (2009) han demostrado teóricamente la evolución espontánea de estos sistemas empíricos en la naturaleza, sin que sean necesarios diseñadores externos que los creen o que los refuercen.

experiencias subjetivas tales como la libertad, equidad, fraternidad, tolerancia, propiedad, justicia, etc., y todos ellos encierran una estrecha relación con sentimientos exogámicos de solidaridad y pertenencia universal (“humanidad” de la ética de los derechos humanos, por ejemplo), que no son incompatibles con sentimientos endogámicos -individualistas o grupalistas- de segregación/indiferencia a los enemigos, a los diferentes o desconocidos (Maryanski y Turner 1992).

El tipo de sociedades en el que el altruismo pasivo es el rasgo predominante en la interacción son las modernas sociedades extensas, caracterizadas por frecuentes encuentros anónimos y con desconocidos, interacciones únicas y carentes de previos vínculos simpatéticos. En la actualidad estos modelos culturales empíricos son las democracias de mercado, sociedades dotadas de avanzadas tecnologías de transporte y comunicación que permiten altas tasas de migración territorial y étnico-cultural. Estos *demes eusociales* son estrictamente diferentes de los superorganismos de himenópteros, así como de los agrupamientos humanos (y homínidos) arcaicos caracterizados por fuertes vínculos de parentesco (Wilson 2005). No obstante, el cerebro biológico y el cerebro cultural pueden estar bien co-adaptados para cooperar con desconocidos si la cultura interiorizada extiende la semejanza fenotípica esencial a la propia naturaleza humana, uno de cuyos rasgos es la diversidad étnica y grupal (véase figura 6.4.) La epigenética de la aproximación tolerante es compartida con otras especies animales, incluso en la expresión física de tal actitud (figura 6.5.)

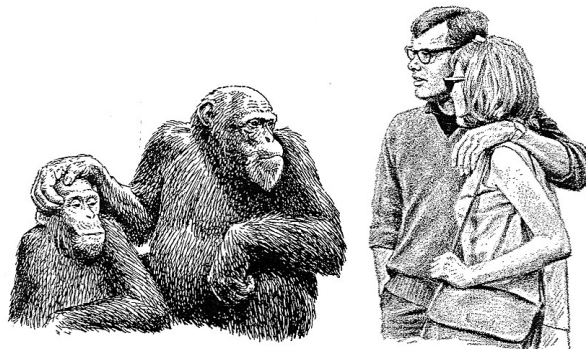
Figura 6.4.- Ilustración de la tolerancia cooperativa entre fenotipos



La fotografía superior representa dos bebés humanos, con un evidente fenotipo diferenciado, abrazados en actitud sosegada. En la infancia los humanos tienden naturalmente a percibir rasgos de humanidad que, combinados con las actitudes de aceptación, juego y mutua protección, producen la experiencia del sosiego y comportamientos cooperativos. La epigénesis de la cooperación se apoya en las ventajas asociadas a la posibilidad de un destino común dentro de relaciones caracterizadas por al menos un altruismo pasivo, o basal.

Fuente: http://www.elconfidencial.com/multimedia/album/mundo/2014-10-13/fotografias-que-retratan-al-ser-humano_236597/#28

Figura 6.5.- Expresión corporal de la actitud tolerante y cooperativa en chimpancés y humanos



El efecto tranquilizador del contacto corporal perdura en los adultos. Izquierda: una hembra de chimpancé, de unos cuatro años, se cobija bajo la mano de un macho. Derecha: una pareja rumana. En Eibl-Eibesfeldt (1994: 122).

Tal y como apostilla Ridley (1996: 6), las culturas no son colecciones aleatorias de hábitos caprichosos, sino expresiones epifenoménicas organizadas de los innatismos de la especie (*human universals*), uno de los cuales es la predisposición a asimilar cultura. En este sentido, no compartimos los supuestos ni las conclusiones de la tesis freudiana del malestar en la cultura -de la sociedad extensa-, presente en trabajos contemporáneos alineados con la tesis de la tragedia de los comunes, y otros que manifiestan su desconfianza acerca del orden y la estabilidad evolutiva de la gran sociedad moderna⁴¹⁰. En ellos se plantea que las sociedades complejas tienen problemas inherentes de supervivencia de las instituciones cooperativas, que al ser características de los micro-órdenes de pequeña escala por ello resultan disfuncionales en la gran sociedad comercial. De ahí se supone el malestar psíquico personal, el conflicto social moderno, y la tendencia a generar tragedias en la gestión de la comunidad y el patrimonio comunitario. Sin embargo la complejidad, sostenemos aquí, puede ser cooperativa porque existen dispositivos neuropsicológicos coherentes y adaptados para la microcooperación (el altruismo personalizado), e igualmente para la macrocooperación (el altruismo anónimo, y el intercambio de mercado), y cada instancia lleva asociadas sus correspondientes instituciones epifenoménicas.

La educación y la interiorización de reglas culturales y morales propias de la civilización humana global es un factor integrador de culturas, y propagador de los rasgos y categorías sociales definitorios de la humanidad de lo humano:

... en cada enunciado del Yo están el cerebro biológico y la cultura social. Cuando el sujeto puede abrir su Nosotros al otro, a sus semejantes, a la vida, el mundo se vuelve rico en humanidad. Por ello, de forma diferente según los individuos y las condiciones, el individuo-sujeto experimenta la autoridad social, la impronta y la norma de una cultura, se forma y vive sin cesar en la dialógica puesta de relieve por Freud entre el Superyó, el Ello pulsional, y el Yo-Yo (*Moi*). (Morin 2004: 86).

En *La teoría de los sentimientos morales* Adam Smith planteó su teoría del espectador imparcial, según la cual uno de los rasgos psicológicos universales de la humanidad de lo humano es el acceso a un estado de consciencia ampliada que permite al yo desapegarse de sus coordenadas históricas e identitarias, y contemplar al otro desde una

⁴¹⁰ Tales como Eibl-Eibesfeldt (1996), algunos de los incluidos en Ostrom, *et. al.* (2002), y hasta el referente teórico de Hayek (1990[1988]), que a su vez está inspirado en Freud (1997[1929]).

posición des-subjetivada propia de la ecuanimidad estoica⁴¹¹. Aquellos aspectos que faciliten la ecuanimidad en la evaluación y valoración de los demás contribuyen a potenciar los sentimientos morales y la humanidad de lo humano. Cualquier creencia en una semejanza pre-cultural de todos los individuos de la especie contribuirá, como mantienen, entre otras, las tradiciones de pensamiento occidental estoica o cristiana, a un mayor cosmopolitismo y una superior apertura a la cooperación con los desconocidos. Y así, a pesar de su posición sobre el malestar humano dentro de la complejidad cosmopolita y tecnológica de la civilización moderna, el propio Freud reconoció en una carta dirigida a Einstein sobre el porqué de la guerra, que:

Todo lo que establezca vínculos afectivos entre los hombres debe actuar contra la guerra. Estos vínculos pueden ser de dos clases. Primero, los lazos que nos ligan a los objetos del amor, aunque desprovistos de fines sexuales (...) La otra forma de vinculación afectiva es la que se realiza por identificación. Cuando establece importantes elementos comunes entre los hombres despierta sentimientos de comunidad, identificaciones. Sobre ellas se funda en gran parte la estructura de la sociedad humana. (Freud 1943[1933(1932)]: 3.213).

El otro puede llegar a ser el “infierno”, como dijera Sartre. Pero también el otro puede ser un cooperador veraz con quien construir benéficos ambientes de cooperación. El infierno no es la alteridad en sí, sino el conflicto real de la lucha violenta entre identidades, así como el desarraigo nihilista de la soledad y el solipsismo.

6.7.- Conclusiones

Aunque la influyente tradición doctrinal del solipsismo atribuye al ser humano una naturaleza desapasionadamente egocéntrica e individualista, sin embargo cualquier observación superficial de la realidad muestra la existencia de relaciones interpersonales altercéntricas, o de sentido grupal y progrupar (*we sense*)⁴¹². A diferencia del modelo de psicología y comportamiento egoísta de la primera caracterización, las relaciones altercéntricas se caracterizan por la cooperación y altruismo, y van desde interacciones diádicas hasta macroprocesos sociales que se expresan en oleadas acción conjunta consistentes con una naturaleza prosocial, e incluso pro-específica, que se cimenta en al

⁴¹¹ Esto tiene su sentido evolutivo, dicho sea de paso, en que para Smith la des-subjetivación, como otras capacidades mentales de intersubjetividad, es una adaptación en beneficio de la especie. Utilizando terminología actual, diríamos que la des-subjetivación interviene modulando la eficacia inclusiva de una población con rasgos homogéneos.

⁴¹² Utilizamos la terminología en la que se apoya Khalil (2004), si bien no necesariamente en el mismo sentido que él lo hace.

menos cierta expresión de amor y respeto al prójimo, a la comunidad y a la humanidad que aquí venimos llamando altruismo pasivo. Esta vendría a ser la esencia terapéutica, y la enseñanza relacional, de todas las religiones, tradiciones sapienciales y sistemas éticos. Es la característica fundamental de la civilización, un reconocimiento del otro como proyección de uno mismo, pero que existe fuera del considerado propio organismo. Sin embargo, los sistemas de valores y la civilización no son creaciones de la mente y del intelecto, sino que la mente y el intelecto humano son creaciones sociales surgidas de la interacción que se produce entre replicadores genéticos, o sea la naturaleza biológica humana, y culturales, o naturaleza cultural humana.

El reconocernos como sujetos auto-conscientes, y el estar en disposición de atribuir esa misma condición a los demás, nos abre al que Thoreau (1999[1854]:11) diría milagro de mirar a través de los ojos del otro. El individuo humano no sólo es un sujeto que concibe a sus semejantes como sujetos, sino que dispone de una intersubjetividad que conecta las subjetividades y las integra en el conocimiento, el sentimiento y la acción. La toma de conciencia de que conocemos y sentimos igual que los demás viene facilitada por la genética de la especie, y reforzada por la epigénesis de estructuras neurales isomórficas, construidas en el proceso de asimilación cultural y de asimilación interpersonal a través de las relaciones. La empatía dispone para la simpatía, motiva intrínsecamente a una apertura prosocial de simpatía alébrica, de modo que, citando a Edgar Morin,

... en cada enunciado del Yo están el cerebro biológico y la cultura social. Cuando el sujeto puede abrir su Nosotros al otro, a sus semejantes, a la vida, el mundo, se vuelven ricos en humanidad. (Morin 2004: 86).

Explican Castro Nogueira, *et. al.* (2003: 184) que el estudio evolutivo de la mente humana muestra, «sin mucho lugar para las dudas, que la concepción del ser humano característica de las ciencias sociales se ha construido sobre presupuestos incompletos». En este capítulo hemos constatado, apoyándonos en una diversidad de trabajos científicos, que la naturaleza humana está adaptada para la sociabilidad y la prosocialidad a través de los procesos psicológicos de intersubjetividad que tienen una causación epigenética. Ello supone que el individuo humano es capaz de concebir utilidades e intereses de otros además de los suyos propios, asimilarlos por

identificación, y actuar en consecuencia de acuerdo a una racionalidad trans-personal o colectiva. La consideración de estas capacidades evolucionadas, y disposiciones conductuales reforzadas por aprendizaje, permite ampliar los horizontes del individualismo metodológico, haciendo posible explicar con mayor riqueza y rigor comportamientos reales de los individuos en la interacción social que la perspectiva egocéntrica explica con dificultad y de una manera forzada y tautológica. Tal es el caso del comportamiento altruista y de la acción conjunta de los colectivos humanos como unidades inteligentes de decisión. La admisión de la racionalidad colectiva (*we rationality*) como hipótesis de trabajo requiere reexaminar la consistencia axiomática y metodológica de las teorías del comportamiento humano construidas exclusivamente sobre la base de la perspectiva egocéntrica de la mente y del comportamiento humano.

Que la especie *Homo sapiens* está capacitada para un modelo evolucionado de intersubjetividad implica que el egoísmo representa sólo una parte de su actividad mental y comportamiento, no la totalidad. La intersubjetividad hace posible la cognición, la emoción, la motivación, y la intención desde una perspectiva plural, que permite la acción concertada con otros individuos sin necesidad de negociación y acuerdos previos. El planteamiento inicial, desarrollo y ejecución de las negociaciones, por otra parte, y en caso de requerirse o realizarse, no son actividades exentas de una intensa intersubjetividad, una capacidad que emplaza al individuo en procesos de comportamiento colectivo a tenor del hecho de que el individuo humano necesariamente desarrolla su ontogénesis en un nicho cognitivo que es la sociedad, donde se lleva a cabo la construcción y la propagación social del conocimiento.

Las personas reales empatizan y simpatizan con sus semejantes, se identifican y vinculan con ellos, y de ahí la perspectiva de acción altruista, la asunción y persecución de objetivos grupales, la acción conjunta, y lo que podemos llamar una racionalidad grupal o plural. Cuando los sujetos están socialmente identificados (aunque se trate de una identificación etnocéntrica y despersonalizada), «actúan ante sí mismos y ante los demás como ejemplares de un grupo, en lugar de como individuos diferenciados» (Caporael, *et. al.* 1989: 693). De no existir un modo de racionalidad conjunta e integradora no sería posible cumplir ordenadamente, y sin continuos conflictos y tragedias de la acción colectiva, las normas sociales elementales, ni proceder

coordinadamente en los procesos de acción conjunta dentro de situaciones cognitivamente complejas. Sería inviable pues la sociedad humana tal y como la conocemos, y por ello la naturaleza humana hubiera evolucionado de otra manera.

En psicología social es conocido que las emociones sociales relacionadas con la inclusión social en grupos con los que el individuo se identifica, favorecen la cohesión por la vía de la percepción de una semejanza subjetivamente acentuada, y expresiones de familiaridad y pertenencia tales como atribuir una superior confianza, honestidad, generosidad e interés recíproco a los miembros del grupo respecto a los *outsiders* (Hoffman 2000: 206 y ss.) Estos procesos emotivos definen el “sesgo de familiaridad”, o “favoritismo inclusivo”, y contribuyen a reforzar la cohesión interna de los grupos a través de la empatía, explicada en el sistema de neuronas espejo, y la vinculación y el compromiso, explicada por las neurohormonas afiliativas. No obstante, la activación de estos sistemas orgánicos es discriminatoria, dadas las limitaciones de la capacidad para empatizar con desconocidos y para desarrollar procesos de acción conjunta con ellos. En efecto, los cerebros simpatizantes son aquellos que manifiestan analogías (e incluso homologías) en sus mapas y modelos emotivo-cognitivos, que son el sustrato de las identificaciones e identidades personales. La semejanza en la estructura encefálica, y su reflejo en los modelos neuropsicológicos de los individuos, es de naturaleza epigenética, o sea, tiene un origen genético y un desarrollo cultural. Como intuyeron tanto Adam Smith como Charles Darwin, la simpatía que conecta las subjetividades de las personas es un instinto, sobre el que proliferan las preferencias de naturaleza cultural.

La identificación con las personas y la identidad grupal explican los comportamientos altruistas humanos sin necesidad de acudir a la teoría de la eficacia inclusiva familiar. Existe la categoría innata del “nosotros”, un instinto tribal, en el diseño funcional del cerebro humano. Sin embargo, el alcance de ese nosotros tiene una explicación y una limitación cultural. La propia naturaleza de los procesos identificatorios, al resultar afectada por la semejanza entre sujeto observador y sujeto observado (isomorfismo genético y neural), asocia la conducta prosocial a coordenadas inclusivas de carácter relacional, en primer lugar (vinculación afiliativa por la acción de

neurohormonas), y de carácter especular (*rapport* simpatizante más rico y preciso por la acción del sistema de neuronas espejo).

Las conductas producto de procesos de identificación tienen un alcance limitado por las propias capacidades cuantitativas del cerebro, y por procesos de modulación asociados al nivel de semejanza neural. La identificación asociada al vínculo y a la empatía, factores explicativos del comportamiento altruista, es superior entre individuos neurológicamente semejantes, y por lo tanto el altruismo resulta de por sí un ejercicio de acciones de eficacia inclusiva entre individuos que portan replicadores comunes. En otro orden de cosas, dado que esta semejanza fenotípica facilita la coordinación y la calidad de la cooperación, el establecimiento de metas comunes, y la ejecución de planes de acción que requieren acción conjunta, la interacción entre individuos vinculados y simpatizantes es de sí un capital social cuyo valor es el resultado de una optimización de eficacia entre las posibles alianzas potenciales (*alliance tracking*). Obviamente, este valor de la cooperación es un valor de supervivencia, que aporta ventajas a los individuos capaces de detectar a sus afines, o de generar entornos institucionales que facilitan la replicación de los rasgos que generan la afinidad.

Una consecuencia importante de lo anterior es que el altruismo humano puede evolucionar a través de estructuras culturales de replicación que producen isomorfismo neural. Esta faceta permite ensanchar el ámbito de la cooperación hasta niveles exponencialmente más amplios que la derivada de la selección de parentesco. La propagación de creencias inclusivas entre individuos capacitados para la vinculación y la simpatía establece preferencias universalistas que contribuyen a la rehumanización de las relaciones humanas. Así, el cerebro biológico y el cerebro cultural pueden estar bien coadaptados para la cooperación exofamiliar y con desconocidos, si las creencias implantadas por la cultura extienden la semejanza fundamental a la propia naturaleza humana, uno de cuyos rasgos es la diversidad cultural adaptativa.

El espectador imparcial que describió Adam Smith como agente des-subjetivado es un producto de la instintiva capacidad humana de la intersubjetividad, reforzada por la asimilación de creencias universalistas sobre la especie. Del éxito en la implantación

de estas creencias, o sus contrarias, se deriva la humanización (benevolencia) o deshumanización (malevolencia) de las relaciones humanas, o lo que es lo mismo, la apertura o restricción de las posibilidades de extraer eficacia de las alianzas potenciales. Como han recomendado todas las tradiciones sapienciales, el conocimiento de sí mismo es la llave de la benevolencia y de la generosidad hacia los demás, ya que el conocimiento de uno mismo permite profundizar el conocimiento de los demás, y a la postre de la especie, y de ahí la comprensión de que la mejor estrategia de supervivencia de un cooperador por naturaleza es la cooperación.

Consecuentemente, la cultura y sus replicadores se muestran decisivos en la evolución de la cooperación humana, dado que incrementan el alcance del grupo dentro del cual se pueden generar dinámicas de selección grupal. He aquí la fortaleza de la sociabilidad cooperativa de la especie, que Spinoza, en una memorable proposición de su *Ética*, afirmara proclamando que «nada es más útil al hombre que el hombre» (Spinoza 2004[1677]): IV, xviii. p. 307-308).

Conclusiones generales

Somos habitados por la vida, por la especie, por nuestros ascendientes, por la cultura, por la sociedad, por las ideas. (Morin 2004: 318).

Cuando aumentamos la iluminación en un área oscura se incrementa también proporcionalmente su contorno de oscuridad. Si utilizamos este fenómeno metafóricamente, puede evaluarse el éxito de esta investigación de acuerdo al grado en que el lector se plantee nuevos interrogantes sobre el tema aquí estudiado.

Hemos intentado esclarecer el complejo y tradicionalmente controvertido estudio del altruismo humano, demostrando que ni psicológica ni materialmente puede admitirse hoy día, sobre la base de las investigaciones realizadas, que exista una completa separación, o dualidad, entre el yo y el otro. Asimismo, la motivación prosocial tiene relevancia empírica y, lo que es más importante, un sentido evolutivo. El altruismo psicológico e intencional es un hecho cierto en la naturaleza humana, existe intrínsecamente, y tiene entidad propia como disposición conductual independiente de otras disposiciones egocentradas y utilitarias. Las adaptaciones que explican la realidad del comportamiento altruista resultan funcionales, han podido evolucionar y, al igual que el egoísmo, tienen valor adaptativo. Por otra parte, los comportamientos prosociales son genuinamente altruistas, con independencia de si quien los realiza extrae o no disfrute u obtiene beneficio derivado de ellos, suponen un coste más o menos elevado para el agente, o se dirijan a individuos emparentados o no, conocidos o desconocidos⁴¹³.

La clave que define el genuino altruismo humano es de tipo intencional y progrupal. Altruista es el comportamiento intencionalmente orientado a mejorar el bienestar de un individuo, un grupo o una población, impulsado por una motivación intrínseca. El agente altruista actúa motivado por la causa última de fortalecer las posibilidades de supervivencia, o sea la eficacia, de otro individuo o agregado de individuos y/o su cultura, entendida esta como una representación simbólica de la identidad del grupo/población. El altruista bien adaptado actúa a favor de una utilidad que le trasciende *qua* individuo, y que integra a otros individuos que conjuntamente constituyen una población de replicadores. El altruista ha de ser capaz de discriminar a

⁴¹³ El disfrute típico del virtuoso entregado a una misión, por ejemplo profesional, que redunde en beneficio de otros, no obsta sin embargo para que tal comportamiento, aun resultando divertido, estimulante o agradable, pueda ser considerado altruista. La intención del científico que dedica su vida a crear una vacuna contra un patógeno es altruista, aunque disfrute intensamente de su trabajo. Agradezco este matiz al profesor Miguel Ángel Quintanilla Fisac, en el IV Encuentro de Estudiantes de Doctorado del Instituto de Estudios de Ciencia y Tecnología, en la Universidad de Salamanca.

los egoístas y defenderse de las estrategias de explotación o predación, ya que de lo contrario su comportamiento resultaría evolutivamente disfuncional. Los replicadores delimitadores de la afinidad pueden ser genéticos y/o culturales, y en ambos casos existe un soporte físico replicable.

La perspectiva poblacional que hemos defendido implica que el altruismo no consiste meramente en una conjetura pensable o un objetivo deseable, como se ha contemplado históricamente desde las visiones escépticas, racionalistas, religiosas o románticas, sino que se trata de un rasgo universal de la especie producto de la evolución, funcional y adaptado para la supervivencia de los replicadores que constituyen la propia especie. Esto significa que sí tiene sentido real admitir que una de las motivaciones humanas es la que podríamos denominar el “bien común”. De esto se deriva que los individuos no actúan siempre y en todo momento motivados por un interés propio estrechamente entendido. Tampoco lo hacen motivados por un altruismo indiscriminado y sin medida. Dependiendo de la composición, estructura e instituciones de la población, los agentes desarrollarán comportamientos adaptativos más o menos inclinados hacia un sentido u otro, admitiendo que existe diversidad de partida en la conformación natural y ambiental de las inclinaciones motivacionales.

Así pues, una conclusión importante a enfatizar derivada de la perspectiva poblacional consiste en que la humanidad como tal no es ni altruista ni egoísta. No existe un individuo o agente prototípico representativo de la especie, sino individuos diversos que forman poblaciones. Estos desarrollan una ontogénesis heterogénea más proclive al altruismo o al egoísmo con mayor o menor intensidad en cada caso, y esa diversidad da lugar a equilibrios poblacionales entre todas las manifestaciones motivacionales y conductuales que interaccionan. La humanidad es, de este modo, una macropoblación con equilibrios evolutivos compatibles con la pluralidad, la adaptabilidad y evolución. La estructura poblacional y las circunstancias ambientales determinan en cada momento el predominio de las gamas de propensiones más o menos egoístas o altruistas. No podemos decir entonces que exista un egoísmo o altruismo innatos definitorios de todos los individuos por igual como naturaleza humana inmutable. Antes bien, *Homo sapiens* ha sobrevivido como especie por ser capaz de

constituir poblaciones con equilibrios variables adaptados a equilibrios circunstanciales, en pos del objetivo último de la supervivencia del grupo y por extensión de la especie. El modo natural de supervivencia y replicación humana es dentro de un grupo/población, por las ventajas adaptativas que esta forma de vida social aporta respecto a la soledad y el aislamiento.

Resulta verosímil pues la tesis de que en los innatismos que permiten la percepción, sentimiento, conocimiento, motivación y acción en *Homo sapiens* existe, entre otras, una disposición filogenética hacia el comportamiento prosocial y proespecífico de tal manera que, cuando el individuo actúa persiguiendo los intereses de otros, p.ej. asistiendo a los necesitados de ayuda, o contribuyendo con externalidades positivas al bien común de la población, coopera sin proponérselo en la mejora y consecución de sus propios intereses.

Esta conclusión supone revisar la validez de la imagen más divulgada de la ciencia económica y otras tradiciones de conocimiento: la hipótesis según la cual cuando el individuo actúa motivado exclusivamente por la consecución de sus intereses como objetivo final de su acción, acaba sin proponérselo beneficiando los intereses generales de la especie. La imagen es cierta en determinados supuestos y circunstancias, según ha sido teóricamente demostrado en el campo de la ciencia económica, pero sin embargo es parcial. Como se puede deducir de la presente tesis, la famosa mano invisible que narrara el economista Adam Smith, como fenómeno complejo de coordinación para el intercambio de bienes y servicios, actúa de acuerdo a procesos relacionales de naturaleza y alcance más diversos y complejos que los normalmente asumidos como explicación de la mano invisible. Somos una especie que juega a que sus individuos y poblaciones coexistan, compitan y cooperen entre sí, en función de equilibrios internos producidos entre una diversidad de patrones de acción agrupables en los conceptos de egoísmo y altruismo. Esto hace posible la división del trabajo y la especialización, así como el intercambio de productos, servicios y conocimiento entre individuos autónomos, y dentro y entre grupos. Parte de la aportación de los individuos a los procesos de especialización y división social del trabajo consiste en la realización de actos altruistas, lo cual es una manifestación de la eusocialidad característica de la

especie (Wilson 2012: 14). Cooperación y competencia son, ambas conductas, y no solamente una de ellas, los fundamentos de la mano invisible real que “dirige” la economía de los recursos dentro de la especie.

Arrow (1994), por ejemplo, señala la condición de vehículos del conocimiento que tienen los individuos, dado que el conocimiento tiene en sí mismo una naturaleza social y de bien público sin la cual no existiría como tal. El conocimiento sólo tiene sentido si se comparte, momento en el que puede pasar a ser cultura siempre que permanezca y se recombine. Si lo logra, pasa a formar parte de la cultura de una población. Si no, se extingue. Arrow deriva de la naturaleza social del conocimiento la existencia de procesos de evolución cultural («*social and historical determinism*», *op. cit.*, p. 4) que superan el alcance del individualismo metodológico como herramienta de análisis de la producción y difusión del conocimiento. Así, cuando se participa en un sistema cultural priman los desequilibrios sobre los equilibrios en los flujos individuales de conocimiento. Las externalidades positivas no internalizadas son connaturales a todo sistema social dinámico e innovador y, como explicó Lévi-Strauss (2008[1949]: 272), el individuo recibe siempre más de lo que aporta en su educación, y simultáneamente aporta más de lo que recibe de su innovación.

Desde la perspectiva de los estudios sociales de ciencia y tecnología, es importante destacar que, además, el comportamiento no sólo de hecho, sino también *intencionalmente* altruista, caracteriza una proporción significativa del quehacer en la producción y difusión del conocimiento. Esto es así en la operativa de su génesis y propagación informal, y también en la que corresponde a las relaciones formales dentro de sistemas institucionales de innovación, ciencia y tecnología. Las relaciones entre maestro-aprendiz, colega-colega, usuario-usuario, o cualquier otra que genere externalidades positivas de difusión informativa, son ejemplos de interacción para la transferencia de conocimiento en que, por su propia naturaleza de legado o patrimonio cultural para el desarrollo de un objetivo común, prima la motivación altruista en un entorno social resultante que es cooperativo (Arena y Conein 2008; Benkler 2012).

Cuando se utiliza un sistema de reglas como estructura necesaria para el desarrollo de un sistema de conocimiento, todos los intervinientes han de acatar esas reglas a fin de poder acceder al conocimiento en sí, y eso mismo implica una actitud cooperativa que permite preservar el espacio en común que se crea. Independientemente de la creación de ese espacio, la propia práctica de la comunicación refuerza el proceso: los individuos que participan en ella han de utilizar un mismo canal (con sus limitaciones inherentes) y eso, que supone de hecho un sacrificio para el individuo (p.ej. ve restringidas sus posibilidades expresivas y creativas), sin embargo, constituye un beneficio desde el punto de vista de la población, que no existiría como tal de no ser así las cosas. El argumento ya fue señalado por Wittgenstein cuando negó teóricamente la posibilidad de un lenguaje privado. Los individuos son a la par artífices y vehículos cooperativos del lenguaje, un lenguaje que en todo caso es un bien común de la población. No obstante, esa condición de vehículos de la transmisión cultural también aporta beneficios (y no solo sacrificios), dado que posibilita el acceso a un acervo de conocimiento que de otra manera ni existiría ni sería asimilable para ellos (Cavalli-Sforza 2010[1996]: 198).

Este argumento aplicado se fundamenta en el hecho de que la regulación de las relaciones entre las partes de un todo reside en un patrón cooperativo inherente. Los sistemas ordenados se caracterizan porque las partes se comportan de forma regular y predecible en orden a la supervivencia del conjunto, de la cual depende directamente la supervivencia de cada una de sus partes (Hayek 2007[1967]: 128). Así pues, la cooperación es una propiedad no sólo de los sistemas de información y conocimiento, sino asimismo de todo sistema físico, lo cual es coherente con la correspondencia e interdependencia de los órdenes materia – vida – mente.

En los seres vivos la cooperación es un producto de la evolución, y en los animales sociales y culturales, como la especie humana, la cooperación es una consecuencia de procesos coevolutivos de replicadores genéticos y culturales. Toda especie es un linaje de replicadores en evolución. En la especie *Homo sapiens* se sabe que sus individuos comparten entre el 99,7% y el 99,9% de la secuencia del genoma (Henderson 2010: 60-61). Como población de replicadores genéticos, la especie

humana constituye un todo ordenado constituido por individuos. Esos individuos actúan, aunque no se percaten de ello, y de acuerdo a una lógica de comportamiento coherente y regulada, en favor de la supervivencia de agregados poblacionales que llegan a incluir el conjunto de la especie. Pero la especie humana, que goza de una sobresaliente homogeneidad genética, es más que una colección ordenada de genes estructurada en macromoléculas independientes de ADN. El 99% del genoma humano es compartido con el chimpancé, y el 97,5% con el ratón (*ibid.*) A pesar de ello, las relaciones de cooperación predominantes en los humanos son intraespecíficas. Esto implica que una fracción muy pequeña de convergencia o divergencia genética (entendida como estructura y pautas de interacción de los genes) puede determinar el sentido cooperativo (o no) de las relaciones entre los organismos. Así pues, es plausible suponer que la cooperación humana es una consecuencia de la propia regulación de los sistemas ordenados de replicadores genéticos y, también, de la existencia de replicadores culturales que, en interacción con los genéticos, se replican cooperativamente posibilitando ventajas evolutivas recíprocas. La cooperación entre dos individuos humanos (que comparten cultura además de genética) es más benéfica para los replicadores compartidos que la que puede surgir de un individuo humano y un ratón, *cæteris paribus* la genética.

Asumido que la cooperación es racional porque surge de adaptaciones eficaces para la supervivencia y propagación no de individuos, sino de rasgos homogéneos dentro de una población, cabe preguntarse acerca de la naturaleza social y grupalista, desde la perspectiva del rasgo, de la racionalidad humana. Desde este punto de vista la racionalidad es una racionalidad colectiva, que conecta, considera e incluye a otros individuos portadores de rasgos homogéneos en los innatismos y programas cognitivos que explican su comportamiento. Así pues, ante la cooperación, es preciso preguntarse el *por qué* y el *para qué* (causación última), y a continuación el *con quién* y *de qué modo* (causación próxima).

Según la teoría del destino común explicada en esta investigación, al ser el resultado probable derivado de no cooperar con nadie $[(x_j), (x_j), \dots]$ inferior, desde la perspectiva de la eficacia del individuo, al de sí hacerlo $[f(x_{ij}), f(x_{ji}), \dots]$, los individuos

desplegarán una racionalidad cooperativa evolucionada siempre que cooperar, ya sea en el marco de una díada (i, j) o un agregado superior (J) produzca que

$$f(x_{i,j}) > x_i$$

Siendo esto así, resulta ecológicamente racional que los cooperadores busquen para interactuar a otros con quienes la cooperación sea más eficiente. Sin embargo, cuando se interactúa en ambientes suficientemente complejos, inciertos y despersonalizados, la búsqueda de los emparejamientos cooperativos no se produce, ni se pueden producir, sobre la base de una supuesta contabilidad racional de beneficios utilitarios. Ante este problema de supervivencia los humanos han desarrollado la capacidad adaptativa de apreciar e identificar determinadas señales, fidedignas y representativas de rasgos explicativos del carácter cooperativo, que les motivan para cooperar con individuos que las expresan, aunque estos sean *a priori* desconocidos. Para el cooperador bien adaptado la decisión no es pues cooperar o no, dado que la cooperación es un supuesto de su naturaleza, sino con quién hacerlo. Las adaptaciones de la coevolución de replicadores biológicos y culturales en buena lógica deben haber producido, aunque sea limitada e imperfectamente, sistemas coherentes de innatismos, emociones, formas de señalización, valores, instituciones y otras normas sociales por los que el individuo, dotado de base altruista, resulte estimulado para actuar de acuerdo a una maximización convencional del tipo buscar al individuo j o al conjunto discreto de individuos $J > 0$ con quien/es cooperar, admitidas las restricciones y limitaciones biológicas y culturales, *hace máxima la diferencia entre eficacia grupal/poblacional, y el coste de oportunidad de cooperar*⁴¹⁴:

$$\text{Max } |[F(X_{i,j}) - x_i]|$$

⁴¹⁴ Los animales, incluidos los humanos, no maximizan intencional ni conscientemente ni eficacia ni ninguna *proxy* asociada, como puede ser la utilidad planteada por la economía. Lo que hacen es que cada individuo, dada su dotación de egoísmo/altruismo (consecuencia de innatismos epigenéticos e *imprintings* culturales aprendidos) responde a los estímulos surgidos de la interacción en un ambiente. Que el resultado sea maximizador o no es una consecuencia de los patrones de comportamiento con los que la población termina respondiendo a los cambios sociales e institucionales (fruto de la coevolución genes-cultura) y medioambientales. En condiciones exigentes de selección cabe esperar que sólo los comportamientos adaptativos puedan sobrevivir.

Dentro de esta vía analítica de la perspectiva poblacional, la semejanza es una variable explicativa de la conformación de estrategias biológicas, así como culturales e institucionales de cooperación. Igualmente, el altruismo ha podido evolucionar porque la especie dispone de adaptaciones para la detección de intenciones e intereses subjetivos, la cognición y la acción sobre la base de una racionalidad plural, y la discriminación y castigo de explotadores y depredadores, todo lo cual disminuye el mencionado coste de oportunidad de cooperar asociado a la no-cooperación activa. Asimismo, existen refuerzos culturales evolucionados (uno es el uso del dinero) para expandir el grado de cooperación y la cuantía de cooperadores con quienes cooperar, interfiriendo en los equilibrios poblacionales de individuos heterogéneos, caracterizados por tener patrones de acción egoísta y altruista. Por ello, las posibilidades de relación que puedan surgir entre cooperadores y no cooperadores en cada equilibrio poblacional configuran y determinan la evolución de la estructura y el orden social, condicionando los entornos sociales, los enclaves y las redes de interacción.

El resultado es que los individuos que tienen un destino común, como es el caso de quienes comparten replicadores biológicos y culturales, manifiestan motivaciones que les conducen a agruparse y comprometerse en procesos de cooperación con aquellos que son semejantes. La cooperación es más eficiente entre ellos porque, al ser semejantes, la intersubjetividad y la identificación son más ricas e intensas, lo cual permite aprovechar, en cada equilibrio poblacional, las sinergias y economías asociadas a la convergencia de sus fenotipos. Esta es la causa última o base evolutiva de las emociones humanas explicativas de la preferencia por lo semejante y el favoritismo inclusivo, y esto último, a la vez, constituye una causa del refuerzo de tales emociones, sustanciado en la actitud de compromiso incondicional producido por la neuroquímica evolucionada de la vinculación interpersonal.

El altruismo humano ha podido evolucionar por selección de grupo porque la cultura favorece la ontogénesis de cerebros semejantes que seleccionan a favor de sus afines con suficiente seguridad, aunque admitiendo el error que permite activar los procesos evolutivos. Para explicar la evolución del altruismo no es preciso pues que exista un gen que codifique el comportamiento altruista. Basta con que existan genes

que codifiquen cerebros capacitados para la neuroplasticidad y la intersubjetividad, y no se requiere que lo hagan de manera perfecta. La cultura posibilita la evolución por selección grupal, así como ampliar la cuantía de población sometida al proceso (y deriva) de la selección de grupo y, además, hacerla consistente con la eficacia mejorada del propio organismo.

Aunque el modelo explicitado es de cooperación entre individuos no emparentados, no obstante, desde la perspectiva de los replicadores explicativos de una naturaleza humana culturalmente adaptable e intersubjetiva, o sea, de un genotipo apto para la cooperación, aparecen fuerzas selectivas de refuerzo a través de selección multinivel. Este patrón evolutivo es posible porque además de individuos egoístas que compiten, explotan y depredan, existen altruistas que cooperan, de cuya interacción surgen dinámicas favorables a la población que son coherentes con su supervivencia individual *qua* cooperadores, y la de aquellos que, sin ser parientes consanguíneos, sí lo son desde la perspectiva de los replicadores que hacen posible la expresión del fenotipo cooperativo y altruista.

En consecuencia, este modelo de selección multinivel exoparental, que admite la interacción con individuos desconocidos que expresan fenotipos resultado de replicadores heredables, y que disponen de capacidades evolucionadas para seleccionar a favor de sus afines, es un modelo general de evolución de la cooperación planteado desde las ciencias sociales, del que la regla de Hamilton puede considerarse un caso particular.

Bibliografía

1. ABULAFIA, David (2009), *El descubrimiento de la humanidad. Encuentros atlánticos en la era de Colón*, Editorial Crítica, Barcelona.
2. ADOLPHS, R.; TRANEL, D.; y DAMASIO, A.R. (1998), “The human amygdala in social judgment”, *Nature*, vol. 393, nº 6.684. pp. 470-474.
3. AKERLOF, George A. (1983), “Loyalty filters”, *American Economic Review*, vol. 73, nº 1. pp. 54-63.
4. AKERLOF, George A., y KRANTON, Rachel E. (2000), “Economics and Identity”, *Quarterly Journal of Economics*, vol. CXV, nº 3. pp. 715-753.
5. AKERLOF, George A., y KRANTON, Rachel E. (2005), “Identity and the economics of organizations”, *Journal of Economic Perspectives*, vol. 19, nº 1. pp. 9-32.
6. AKERLOF, George A.; y KRANTON, Rachel E. (2008), “Identity, supervision, and work groups”, *American Economic Review*, vol. 98, nº 2. pp. 212-217.
7. AKERLOF, George A.; y YELLEN, Janet L. (1986) (eds.), *Efficiency wage models and the labor market*, Cambridge University Press, Cambridge & New York.
8. AKNIN, Lara B., DUNN, Elizabeth W., WHILLANS, Ashley V., GRANT, Adam M., NORTON, Michael I., (2013), “Making a difference matters: Impact unlocks the emotional benefits of prosocial spending”, *Journal of Economic Behavior and Organization*, vol. 88, nº 1. pp. 90-95.
9. AKNIN, Lara B., SANDSTROM, Gilliam M., DUNN, Elizabeth, W., y NORTON, Michael I. (2011), “It's the recipient that counts: Spending money on strong social ties leads to greater happiness than spending on weak social ties”, *PloS ONE*, vol. 6, nº 2. e17018.
10. ALCHIAN, Armen A. (1950), “Uncertainty, evolution and economic theory”, *Journal of Political Economy*, vol. 58, nº 3. pp. 211-221.
11. ALCHIAN, Armen A., y DEMSETZ, Harold (1994[1972]), “Producción, costes de información y organización económica”, en Putterman (1994). pp. 141-170.
12. ALEXANDER, Richard D. (1971), “The search for an evolutionary philosophy”, *Proceedings of the Royal Society Victoria of Australia*, vol. 84. pp. 99-120.
13. ALEXANDER, Richard D. (1974), “The evolution of social behavior”, *Annual Review of Ecology Systematics*, vol. 5. pp. 325-383.
14. ALEXANDER, Richard D. (1975), “The search for a general theory of behavior”, *Behavioral Sciences*, vol. 20, nº1. pp. 77-100.

15. ALEXANDER, Richard D. (1987), *The biology of moral systems*, Aldine de Gruyter, New York.
16. ALGER, Ingela, y WEIBULL, Jörgen W. (2012), “A generalization of Hamilton's rule – Love others how much?”, *Journal of Theoretical Biology*, vol. 299, nº 1. pp. 42-54.
17. ALGOE, Sara B., y HAIDT, Jonathan (2009), “Witnessing excellence in action: the 'other-praising' emotions of elevation, gratitude, and admiration”, *Journal of Positive Psychology*, vol. 4, nº 2. pp. 105-127.
18. AMATO, Paul R. (1990), “Personality and social network involvement as predictor of helping behavior in everyday life”, *Social Psychology Quarterly*, vol. 53, nº 1. pp. 31-43.
19. ANDERSON, Christopher M.; y PUTTERMAN, Louis (2006), “Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism”, *Games and Economic Behavior*, vol. 54, nº 1. pp. 1-24.
20. ANDREONI, James (1988), “Why free ride? Strategies and learning in public good experiments”, *Journal of Public Economics*, vol. 37, nº 2. pp. 291-304.
21. ANDREONI, James (1989), “Giving with impure altruism: Applications to charity and Ricardian equivalence”, *Journal of Political Economy*, vol. 97, nº 6. pp. 1.447-1.458.
22. ANDREONI, James (1990), “Impure altruism and donations to public goods: A theory of warm-glow giving”, *Economic Journal*, vol. 100, nº 401. pp. 464-477.
23. ANDREONI, James (1993), “An experimental test of the public-goods crowding-out hypothesis”, *American Economic Review*, vol. 83, nº 5. pp. 1.317-1.327.
24. ANDREONI, James, y CROSON, Rachel (2008), “Partners versus strangers: Random rematching in public goods experiments”, en PLOTT, Charles R., y SMITH, Vernon L. (eds.), *Handbook of Experimental Economics Results*, vol. 1, North-Holland, Amsterdam. pp. 776-783.
25. ANDREONI, James, y RAO, Justin M. (2011), “The power of asking: How communication affects selfishness, empathy, and altruism”, *Journal of Public Economics*, vol. 95. nº 7-8. pp. 513-520.
26. ANDREONI, James, y VERTERLUND, Lise (2001), “Which is the fair sex? Gender differences in altruism”, *Quarterly Journal of Economics*, vol. 116, nº 1. pp. 293-312.
27. ANDREONI, James; HARBAUGH, William T.; y VESTERLUND, Lise (2008), “Altruism in experiments”, en DURLAUF, Steven N., y BLUME, Lawrence E., (eds.), *The New Palgrave Dictionary of Economics*, 2nd edition, Palgrave Macmillan. Edición en hipertexto:

http://www.dictionaryofeconomics.com/article?id=pde2008_A000240
doi:10.1057/9780230226203.0035

28. ANDREONI, James; y MILLER, John H. (1993), "Rational cooperation in the finitely repeated Prisoner's Dilemma: Experimental Evidence", *The Economic Journal*, vol. 103, n° 418. pp. 570-585.
29. AOKI, Kenichi (1986), "A stochastic model of gene-culture coevolution suggested by the 'culture historical hypothesis' for the evolution of adult lactose absorption in humans", *Proceedings of the National Academy of Sciences*, vol. 83, n° 5. pp. 2.929-2.933.
30. APPIAH, Kwame A. (2007), *La ética de la identidad*, Katz Editores, Buenos Aires.
31. ARENA, Richard; y CONEIN, Bernard (2008), "On virtual communities: Individual motivations, reciprocity and we-rationality", *International Review of Economics*, vol. 55, n° 1-2. pp. 185-208.
32. ARENDT, Hannah (2001[1929]), *El concepto de amor en san Agustín*, Ediciones Encuentro, Madrid.
33. ARENDT, Hannah (2002[1978]), *La vida del espíritu*, Editorial Paidós, Barcelona.
34. ARENDT, Hannah (2003[1982]), *Conferencias sobre la filosofía política de Kant*, introducción y edición a cargo de Ronald Beiner, Editorial Paidós, Barcelona.
35. ARISTÓTELES (1994), *Política*, traducción, prólogo y notas de Carlos García Gual y Aurelio Pérez Jiménez, Alianza Editorial, Madrid.
36. ARISTÓTELES (2002), *Ética Eudemia*; traducción, introducción y notas de Carlos Megino Rodríguez, Alianza Editorial, Madrid.
37. ARROW, Kenneth J. (1974), *The limits of organization*, W.W. Norton & Company, New York.
38. ARROW, Kenneth J. (1994), "Methodological individualism and social knowledge", *American Economic Review*, vol. 84, n° 2. pp. 1-9.
39. ARSUAGA, Juan L. (2001), "El misterio del neandertal", *El País Semanal*, n° 1.316, 16 de diciembre. p. 58.
40. ASHFORTH, Blake E., y MAEL, Fred A. (1989), "Social identity theory and the organization", *Academy of Management Review*, vol. 14, n° 1. pp. 20-39.
41. AUMANN, Robert J. (2005), "War and Peace", Nobel Prize Lecture, december 8th, 2005. pp. 350-358. Edición en hipertexto: http://nobelprize.org/nobel_prizes/economics/laureates/2005/aumann-lecture.pdf
42. AUSTIN, James H. (2006), *Zen-brain reflections*, The MIT Press, Cambridge (Mass.) and London.

43. AXELROD, Robert (1981), "The emergence of cooperation among egoists", *American Political Science Review*, vol. 75, nº 2. pp. 306-318.
44. AXELROD, Robert (1996[1984]), *La evolución de la cooperación. El dilema del prisionero y la teoría de juegos* [1984], versión española de Luis Bou, Alianza Editorial, Madrid.
45. AXELROD, Robert (1997), *The complexity of cooperation: Agent-based models of competition and collaboration*, Princeton University Press, Princeton (NJ).
46. AXELROD, Robert, y DION, Douglas (1988), "The further evolution of cooperation", *Science*, vol. 242, n1 4.884. pp. 1.385-1.390.
47. AXELROD, Robert, y HAMILTON, William D. (1981), "The evolution of cooperation", *Science*, vol. 211, nº 4.489. pp. 1.390-1.396.
48. AYALA, Francisco J. (1994), *La naturaleza inacabada. Ensayos en torno a la evolución*, Salvat Editores, Barcelona.
49. AYER, Alfred J. (1979[1954]), "Libertad y necesidad", en *Ensayos filosóficos*, traducción de Francisco Béjar, Ariel, Barcelona. Cap. 12.
50. BACHARACH, Michael (1999), "Interactive team reasoning: A contribution to the theory of cooperation", *Research in Economics*, vol. 53, nº 2. p. 117-147.
51. BACK, István, y FLACHE, Andreas (2006), "The viability of cooperation based on interpersonal commitment", *Journal of Artificial Societies and Social Simulation*, vol. 9, nº 1. Edición en hipertexto: <http://jasss.soc.surrey.ac.uk/9/1/12.html>
52. BACON, Francis (1985[1620]), *Novum Organum. Aforismos sobre la interpretación de la naturaleza y el reino del hombre*, traducción y prólogo de Texeira Bastos, Editorial Orbis, Barcelona.
53. BAGUS, Philipp (2004), "La tragedia de los bienes comunales y la escuela austriaca: Hardin, Hoppe, Huerta de Soto y Mises", *Procesos de Mercado: Revista Europea de Economía Política*, vol. I, nº 2, otoño. pp. 125-139.
54. BALLIET, Daniel (2010), "Communication and cooperation in social dilemmas: A meta-analytic review", *Journal of Conflict Resolution*, vol. 54, nº 1. pp. 39-57.
55. BANCO MUNDIAL (2015), *World Development Report 2015. Mind, society and behavior*, The World Bank, Washington.
56. BANDIERA, Oriana; BARANKAY, Iwan; y RASUL, Imran (2005), "Cooperation in collective action", *Economics of Transition*, vol. 13, nº 3. pp. 473-498.

57. BANKS, William P., e ISHAM, Eve A. (2009), "We infer rather than perceive the moment we decide to act", *Psychological Science*, vol. 20, nº 1. pp. 17-21.
58. BARCLAY, Pat (2004), "Trustworthiness and competitive altruism can also solve the 'tragedy of the commons'" *Evolution and Human Behavior*, vol. 25, nº 4. pp. 209-220.
59. BARKOW, Jerome H; COSMIDES, Leda; y TOOBY, John (eds.) (1992), *The adapted mind: Evolutionary psychology and the generation of culture*, Oxford University Press, Oxford & New York.
60. BARKOW, Jerome H. (1978), "Social norms, the self, and sociobiology: Building on the ideas of A.I. Hallowell", *Current Anthropology*, vol. 19, nº 1. pp. 99-103.
61. BARNARD, Chester I. (1968[1938]), *The functions of the executive*, Harvard University Press, Cambridge, Massachussets.
62. BARON-COHEN, Simon (1995), *Mindblindness. An essay on autism and theory of mind*. MIT Press, Cambridge, Massachussets.
63. BARRAZA, Jorge A., y ZAK, Paul J. (2009), "Empathy toward strangers triggers oxytocin release and subsequent generosity", *Annals of the New York Academy of Sciences*, vol. 1.167. pp. 182-189.
64. BARSALOU, Lawrence W. (1999), "Perceptual symbol systems", *Behavioral and Brain Sciences*, vol. 22, nº 4. pp. 577-660.
65. BARSALOU, Lawrence W., SIMMONS, W. Kyle, BARBEY, Aron K., y WILSON, Christine D. (2003), "Grounding conceptual knowledge in modality-specific systems", *TRENDS in Cognitive Sciences*, vol. 7, nº 2. pp. 84-91.
66. BARTELS, Andreas, y ZEKI, Semir (2004), "The neural correlates of maternal and romantic love", *Neuroimage*, vol. 24, nº 3. pp. 1.155-1.166.
67. BARTRA, Roger (2006), *Antropología del cerebro. La conciencia y los sistemas simbólicos*, Editorial Pre-Textos, Valencia.
68. BATSON, C. Daniel (1991), *The altruism question: Toward a social psychological answer*, Lawrence Erlbaum Associates, Hillsdale (NJ).
69. BATSON, C. Daniel (1992), "Experimental tests for the existence of altruism", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1992, nº 2. pp. 69-78.
70. BATSON, C. Daniel, LISHNER, D. A., COOK, J., y SAWYER, S. (2005), "Similarity and nurturance: Two possible sources of empathy for strangers", *Basic and Applied Social Psychology*, vol. 27, nº 1. pp. 15-25.
71. BATSON, C. Daniel; FULTZ, Jim; SCHOENRADE, Patricia A.; y PADUANO, Alan (1987), "Critical self-reflection and self-perceived

- altruism: When self-reward fails”, *Journal of Personality and Social Psychology*, vol. 53, nº 3. pp. 594-602.
72. BATSON, C. D., PATE, S., LAWLESS, H., SPARKMAN, P., LAMBERS, S., WORMAN, B. (1979), “Helping under conditions of common threat: increased 'we-feeling' or ensuring reciprocity”, *Social Psychological Quarterly*, vol. 42, nº 4. pp. 410-414.
 73. BAUMEISTER, Roy F., y LEARY, Mark (1995), “The need to belong: Desire for interpersonal attachments as a fundamental human motivation”, *Psychological Bulletin*, vol. 117, nº 3. pp. 497-529.
 74. BEAOUREGARD, Mario, y VINCENT, Paquette (2006), “Neural correlates of a mystical experience in carmelite nuns”, *Neuroscience Letters*, vol. 405, nº 3. pp. 186-190.
 75. BECHARA, Antoine (2005), “Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective”, *Nature Neuroscience*, vol. 8, nº 11. pp. 1.458-1.463.
 76. BECKER, Gary S. (1981), “Altruism in the family and selfishness in the market place”, *Economica*, vol. 48, nº 189. pp. 1-15.
 77. BECKER, Gary S. (1987), *Tratado sobre la familia*, primera edición en español, Alianza Editorial, Madrid.
 78. BECKER, Gary S. (1996), “Spouses and beggars: love and sympathy”, en *Accounting for tastes*, Harvard University Press, Cambridge, Massachusetts. pp. 231-237.
 79. BECKER, Gary S. (1997[1962]), “The irrational behavior and the economic theory”, *Journal of Political Economy*, vol. 70, nº 1. Traducido al castellano en FEBRERO, Ramón, y SCHWARTZ, Pedro (1997), *La esencia de Becker*, Editorial Ariel, Barcelona. pp. 59-73.
 80. BECKER, Gary S. (1997[1971]), *The economics of discrimination*, 2ª edición, University of Chicago Press, Chicago. Extractado y traducido al castellano: “Las fuerzas determinantes de la discriminación en el mercado”, en FEBRERO, Ramón, y SCHWARTZ, Pedro (1997), *La esencia de Becker*, Editorial Ariel, Barcelona. pp. 401-411.
 81. BECKER, Gary S. (1997[1974]), “A theory of social interactions”, *Journal of Political Economy*, vol. 82, nº 6. Traducido al castellano en FEBRERO, Ramón, y SCHWARTZ, Pedro (1997), *La esencia de Becker*, Editorial Ariel, Barcelona. pp. 177-205.
 82. BECKER, Gary S. (1997[1974a]), “A theory of marriage II”, *Journal of Political Economy*, vol. 82, nº 2. Traducido al castellano en FEBRERO, Ramón, y SCHWARTZ, Pedro (1997), *La esencia de Becker*, Editorial Ariel, Barcelona. pp. 319-334.
 83. BECKER, Gary S. (1997[1976a]), “Altruism, egoism, and genetical fitness: Economics and Sociobiology”, *Journal of Economic Literature*,

- vol. 14, nº 3. Traducido al castellano en FEBRERO, Ramón, y SCHWARTZ, Pedro (1997), *La esencia de Becker*, Editorial Ariel, Barcelona. pp. 335-347.
84. BECKER, Gary S. (1997[1976b]), *The economic approach to human behavior*, University of Chicago Press, Chicago. Traducción al castellano del primer capítulo del libro en FEBRERO, Ramón, y SCHWARTZ, Pedro (1997), *La esencia de Becker*, Editorial Ariel, Barcelona. pp. 47-58.
 85. BECKER, Gary S.; y MURPHY, Kevin M. (2000), *Social economics. Market behavior in a social environment*, The Belknap Press of Harvard University Press, Cambridge (Massachusetts) & London.
 86. BEN-NER, Avner; y PUTTERMAN, Louis (1998), “Values and institutions in economic analysis”, en BEN-NER, A.; y PUTTERMAN, L. (comps.), *Economics, values and organization*, Cambridge University Press, New York. pp. 3-72.
 87. BEN-PORATH, Yoram (1980), “The F-Connection: Families, Friends, and Firms and the Organization of Exchange”, *Population and Development Review*, vol. 6, nº 1. pp. 1-30.
 88. BENKLER, Yochai (2012), *El pingüino y el Leviatán*, traducción de Jorge Paredes, Ediciones Deusto, Barcelona.
 89. BENTHAM, Jeremy (2000[1781]), *An introduction to the principles of morals and legislation*, Batoche Books, Kitchener, Ontario.
 90. BEREZKEI, Tamas, BIRKAR, Bela, y KEREKES, Zsuzsanna (2010), “Altruism toward strangers in need: Costly signaling in an industrial society”, *Evolution and Human Behavior*, vol. 31, nº 1. pp. 95-103.
 91. BERG, Joyce; DICKHAUT, John; y McCABE, Kevin (1995), “Trust, reciprocity, and social history”, *Games and Economic Behavior*, vol. 10, nº 1. pp. 122-142.
 92. BERGER, Peter L.; y LUCKMANN, Thomas (1986[1966]), *La construcción social de la realidad*, traducción de Silvia Zulueta, H. F. Martínez de Murgía, Madrid.
 93. BERGSTROM, Theodore C. (2002), “Evolution of social behavior: Individual and group selection”, *Journal of Economic Perspectives*, vol. 16, nº 2. pp. 67-88.
 94. BERLIN, Isaiah (1978), “Two concepts of liberty”, en QUINTON, Anthony (ed.), *Political Philosophy*, Oxford University Press, Oxford. pp. 141-152.
 95. BINMORE, Ken (1994), *Playing fair: Game theory and the social contract I*, MIT Press, Cambridge, Massachusetts.

96. BINMORE, Ken (1998), *Just playing: Game theory and the social contract II*, MIT Press, Cambridge, Massachusetts.
97. BINMORE, Ken (2005), *Natural Justice*, Oxford University Press, Oxford.
98. BINMORE, Ken (2006a), “The origins of fair play”, en *Papers on Economics and Evolution*, Evolutionary Economics Group, Max Planck Institute of Economics, nº 0614/2006, Jena (Germany). Edición en hipertexto: <https://papers.econ.mpg.de/evo/discussionpapers/2006-14.pdf>
99. BINMORE, Ken (2006b), “Why do people cooperate?”, *Politics, Philosophy & Economics*, vol. 5, nº 1. pp. 81-96.
100. BINMORE, Ken (2007), “Economic man -or straw man?”, Centre for Economic Learning and Social Evolution, Paper nº 262/2007. Edición en hipertexto: <http://else.econ.ucl.ac.uk/papers/uploaded/262.pdf>
101. BINMORE, Ken (2009), *La teoría de juegos. Una breve introducción*, traducción de José Ventura López, Alianza Editorial, Madrid.
102. BISIN, Alberto, y VERDIER, Thierry (2001), “The economics of cultural transmission and the dynamics of preferences”, *Journal of Economic Theory*, vol. 97, nº 2. pp. 298-319.
103. BISIN, Alberto, y VERDIER, Thierry (2008), “Cultural transmission”, en DURLAUF, Steven N., y BLUME, Lawrence E., (eds.), *The New Palgrave Dictionary of Economics*, 2nd edition, Palgrave Macmillan. [The New Palgrave Dictionary of Economics Online, Palgrave Macmillan, 13 January 2010. Edición en hipertexto: http://www.dictionaryofeconomics.com/article?id=pde2008_C000549 doi:10.1057/9780230226203.0345]
104. BLACKMORE, Susan (1999), *The meme machine*, Oxford University Press, Oxford & New York.
105. BLACKMORE, Susan (2009), *Conversations on consciousness*, Oxford University Press, Oxford.
106. BLAFFER, Sarah (2009), *Mothers and others. The evolutionary origins of mutual understanding*, The Belknap Press of Harvard University Press, Cambridge (Mass.) & London.
107. BOCHET, Olivier; PAGE, Talbot; y PUTTERMAN, Louis (2006), “Communication and punishment in voluntary contribution experiments”, *Journal of Economic Behavior and Organization*, vol. 60, nº 1. pp. 11-26.
108. BOEHM, Christopher (1993), “Egalitarian society and reverse dominance hierarchy”, *Current Anthropology*, vol. 34, nº 3. pp. 227-254.

109. BOHNET, Iris, y FREY, Bruno S. (1999a), "The sound of silence in prisoner's dilemma and dictator games", *Journal of Economic Behavior and Organization*, vol. 38, n° 1. pp. 43-57.
110. BOHNET, Iris, y FREY, Bruno S. (1999b), "Social distance and other regarding behavior in dictator games: Comment", *American Economic Review*, vol. 89, n° 1. pp. 335-339.
111. BOLAND, Lawrence A. (1982), *The foundations of economic method*, Allen & Unwin, London.
112. BOLTON, Gary E., KATOL, Elena, Y OCKENFELS, Axel (2005), "Cooperation among strangers with limited information about reputation", *Journal of Publica Economics*, vol. 89, n° 8. pp. 1.457-1.468.
113. BOLTON, Gary E., y OCKENFELS, Axel (2000), "ERC: A theory of equity, reciprocity, and competition", *American Economic Review*, vol. 90, n° 1. pp. 166-193.
114. BOND, Meg A.; y PYLE, Jean L. (1998), "Diversity dilemmas at work", *Journal of Management Inquiry*, vol. 7, n° 3. pp. 252-269.
115. BÖRGERS, Tilman, y SARIN, Rajiv (1997), "Learning through reinforcement and replicator dynamics", *Journal of Economic Theory*, vol. 77, n° 1.
116. BORJIGIN, Jimo, UnCHEOL, Lee, LIU, Tiecheng, PAL, Dinesh, HUFF, Sean, KLARR, Daniel, SLOBODA, Jennifer, HERNANDEZ, Jason, WANG, Michael M., y MASHOUR, George A. (2013), "Surge of neurophysiological coherence and connectivity in the dying brain", *Proceedings of the National Academy of Sciences*, vol. 110, n° 35. pp. 14.432-14.437.
117. BOULDING, Kenneth E. (1971b), "The economics of human conflict", en *Collected Papers*, vol. II, Colorado Associated University Press, Boulder, Colorado. pp. 325-344.
118. BOULDING, Kenneth E. (1975[1953]), *La revolución organizativa. Un estudio sobre la ética de la organización económica*, Monte Avila Editores, Caracas.
119. BOULDING, Kenneth E. (1976), *La economía del amor y del temor. Una introducción a la economía de las donaciones*, Alianza Editorial, Madrid.
120. BOWLES, Samuel (2003), *Microeconomics: Behavior, institutions and evolution*, Princeton University Press, Princeton.
121. BOWLES, Samuel (2006), "Group competition, reproductive leveling, and the evolution of human altruism", *Science*, vol. 314, n° 5.805. pp. 1.569-1.572.

122. BOWLES, Samuel (2008), "Conflict: Altruism's midwife", *Nature*, vol. 456, n° 7.220. pp. 326-327.
123. BOWLES, Samuel; BOYD, Robert; CAMERER, Colin F.; FEHR, Ernst; GINTIS, Herbert; y HENRICH, Joseph (eds.) (2004), *Foundations of human sociality*, Oxford University Press, Oxford.
124. BOWLES, Samuel; BOYD, Robert; FEHR, Ernst; y GINTIS, Herbert (1997), "Homo reciprocans: A research initiative on the origins, dimensions, and policy implications of reciprocal fairness", *Advances in complex systems*, vol. 4, n° 1. pp. 1-30.
125. BOWLES, Samuel; y GINTIS, Herbert (1998), "The moral economy of communities: Structured populations and the evolution of pro-social norms", *Evolution and Human Behavior*, vol. 19, n° 1. pp. 3-25.
126. BOWLES, Samuel; y GINTIS, Herbert (2001), "Social capital and community governance", *Economic Journal*, vol. 112, n° 483. pp. 419-436.
127. BOWLES, Samuel; y GINTIS, Herbert (2002), "Prosocial emotions", Santa Fe Institute Working Paper, n° 02-07-028. Edición en hipertexto: <http://www.santafe.edu/research/publications/wpabstract/200207028>
128. BOWLES, Samuel; y GINTIS, Herbert (2003), "Origins of human cooperation", en HAMMERSTEIN, Peter (ed.), *The Genetic and cultural evolution of cooperation*, MIT Press, Cambridge, Massachusetts. pp. 429-443.
129. BOWLES, Samuel; y GINTIS, Herbert (2004), "The evolution of strong reciprocity: Cooperation in heterogeneous populations", *Theoretical Population Biology*, vol. 65, n° 1. pp. 17-28.
130. BOWLES, Samuel; y GINTIS, Herbert (2008), "Cooperation", en DURLAUF, Steven N., y BLUME, Lawrence E., (eds.), *The New Palgrave Dictionary of Economics*, 2nd edition, Palgrave Macmillan. [The New Palgrave Dictionary of Economics Online, Palgrave Macmillan, 13 January 2010. Edición en hipertexto: http://www.dictionaryofeconomics.com/article?id=pde2008_C000597 doi:10.1057/9780230226203.0317]
131. BOWLES, Samuel; y GINTIS, Herbert (2011), *A cooperative species. Human cooperation and its evolution*, Princeton University Press, Princeton (NJ).
132. BOYD, Robert (1992), "The evolution of reciprocity when conditions vary", en HANCOURT, Alexander, y DeWAAL, Frans B. M. (eds.), *Coalitions and alliances in humans and other animals*, Oxford University Press, Oxford. pp. 473-489.
133. BOYD, Robert, BOGERHOFF-MULDER, Monique, DURHAM, William, y RICHERSON, Peter J. (1997), "Are cultural phylogenies

- possible?”, en WEINGART, Peter, RICHERSON, Peter J., MITCHELL, Sandra D., y MASSEN, Sabine (eds.), *Human by nature. Between biology and the social sciences*, Lawrence Erlbaum Associates, London & Mahwah (NJ). pp. 355-386.
134. BOYD, Robert, y RICHERSON, Peter J. (1985), *Culture and the evolutionary process*, University of Chicago Press, Chicago.
 135. BOYD, Robert, y RICHERSON, Peter J. (1987), “The evolution of ethnic markers”, *Cultural Anthropology*, vol. 2, nº 1. pp. 65-79.
 136. BOYD, Robert, y RICHERSON, Peter J. (1993), “Rationality, imitation, and tradition”, en DAY, R.; y CHEN, P. (eds.), *Nonlinear dynamics and evolutionary economics*, Oxford University Press, New York. pp. 131-149.
 137. BOYD, Robert, y RICHERSON, Peter J. (2001), “Norms and bounded rationality”, en GIGERENZER, Gerd; y SELTEN, Richard (eds.), *Bounded rationality: The adaptive toolbox*, MIT Press, Cambridge. pp. 281-296.
 138. BOYD, Robert, y RICHERSON, Peter J. (2002), “Group beneficial norms can spread rapidly in a structured population”, *Journal of Theoretical Biology*, vol. 215, nº 3. pp. 287-296.
 139. BOYD, Robert, y RICHERSON, Peter J. (2008), “Gene-culture coevolution and the evolution of social institutions”, en ENGEL, Christoph, y SINGER, Wolf (eds.), *Decision making, the human mind, and implications for institutions*, MIT Press, Cambridge, Massachusetts. pp. 305-323.
 140. BOYD, Robert, y SILK, Joan B. (2004), *Cómo evolucionaron los humanos*, 2ª edición, edición española a cargo de Jaume Bertranpetir y David Comas, traducción de David Comas, *et. al.*, Editorial Ariel, Barcelona.
 141. BOYD, Robert; GINTIS, Herbert; BOWLES, Samuel; y RICHERSON, Peter J. (2003), “The evolution of altruistic punishment”, *Proceedings of the National Academy of Sciences, USA*, vol. 100, nº 6. pp. 3.531-3.535.
 142. BOYD, Robert; RICHERSON, P.J.; BORGERHOFF, M., y DURHAM, W.H. (1997), “Are cultural phylogenies possible?”, en WEINGART, Peter; RICHERSON, Peter J.; MITCHELL, Sandra D.; y MAASEN, Sabine (eds.), *Human by nature. Between biology and the social sciences*, Lawrence Erlbaum Associates Publishers, Mahwah (New Jersey) & London. pp. 355-386.
 143. BOYD, Robert; y MATHEW, Sarah (2007), “A narrow road to cooperation”, *Science*, vol. 316, nº 5.833. pp. 1.858-1.859.

144. BRANDT, Hannelore; y SIGMUND, Karl (2004), "The logic of reprobation: Assessment and action rules for indirect reciprocity", *Journal of Theoretical Biology*, vol. 231, nº 4. pp. 475-486.
145. BRAÑAS-GARZA, Pablo (2006), "Poverty in dictator games: Awakening solidarity", *Journal of Economic Behavior and Organization*, vol. 60, nº 3. pp. 306-320.
146. BRAÑAS-GARZA, Pablo (2007), "Promoting helping behavior with framing in dictator games", *Journal of Economic Psychology*, vol. 28, nº 4. pp. 477-486.
147. BRAÑAS-GARZA, Pablo; DURÁN, María; y ESPINOSA, María Paz (2009), "The role of personal involvement and responsibility in unfair decisions", *Rationality and Society*, vol. 21, nº 2. pp. 225-248.
148. BREFCZYNSKI-LEWIS, J.A., LUTZ, A., SCHAEFER, H.S., LEVINSON, D.B., y DAVIDSON, R.J. (2007), "Neural correlates of attentional expertise in long-term meditation practitioners", *Proceedings of the National Academy of Sciences*, vol. 104, nº 27. pp. 11.483-11.488.
149. BREWER, Marilyn B.; y KRAMER, Roderick M. (1986), "Choice behavior in social dilemmas. Effects of social identity, group size, and decision framing", *Journal of Personality and Social Psychology*, vol. 50, nº 3. pp. 543-549.
150. BRONCANO, Fernando (2004), "Capacidades metarrepresentacionales y conducta simbólica", *Estudios de Psicología*, vol. 25, nº 2. pp. 183-203.
151. BROSIG, Jeanette ; OCKENFELS, Axel ; y WEIMANN, Joachim (2003), "The effect of communication media on cooperation", *German Economic Review*, vol. 4, nº 2. pp. 217-241.
152. BROSNAN, Sarah F. (2008), "Fairness and other-regarding preferences in nonhuman primates", en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 77-105.
153. BROSNAN, Sarah F., y De WAAL, Frans B. M. (2003), "Monkeys reject unequal pay", *Nature*, vol. 425, nº 6.955. pp. 297-299.
154. BROTHERS, Leslie (1989), "A biological perspective on empathy", *American Journal of Psychiatry*, vol. 146, nº 1. pp. 10-19.
155. BROWN, S. L., NESSE, R. M., VINOKUR, A. D., y SMITH, D. M. (2003), "Providing social support may be more beneficial than receiving it: Results from a prospective study of mortality", *Psychological Science*, vol. 14, nº 4. pp. 320-327.
156. BROWN, William M., y MOORE, Chris (2000), "Is prospective altruist-detection an evolved solution to the adaptive problem of subtle cheating

- in cooperative ventures? Supportive evidence using Wason selection task”, *Evolution and Human Behavior*, vol. 21, nº 1. pp. 25-37.
157. BROWNING, Christopher R. (2002), *Aquellos hombres grises. El Batallón 101 y la Solución Final en Polonia*, traducción de Montse Batista, Edhasa, Barcelona.
 158. BRUHN, John G. (2005), “Culture-brain interactions”, *Integrative Psychological and Behavioral Science*, vol. 40, nº 4. pp. 243-254.
 159. BRUNEAU, Emile G., DUFOUR, Nicholas, y SAXE, Rebecca (2012), “Social cognition in members of conflict groups: Behavioral and neural responses in Arabs, Israelis, and South Americans to each other's misfortunes”, *Philosophical Transactions of the Royal Society, Biological Sciences*, vol. 367, nº 1.589. pp. 717-730.
 160. BRUNEAU, Emile G., y SAXE, Rebecca (2012), “The power of being heard: The benefits of perspective giving in the context of intergroup conflict”, *Journal of Experimental Social Psychology*, vol. 48, nº 4. pp. 855-866.
 161. BUCHAN, J. Ross, YOON, Je-Hyun, y PARKER, Roy (2011), “Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*”, *Journal of Cell Science*, vol. 124, nº 2. pp. 228-239.
 162. BUCHANAN, James M. (1965a), “An economic theory of clubs”, *Economica*, vol. 32, nº 125. pp. 1-14.
 163. BUCHANAN, James M. (1965b), “Ethical rules, expected values, and large numbers”, *Ethics*, vol. 76, nº 1. pp. 1-13.
 164. BUCHANAN, James M. (1978), “Markets, States and the Extent of Morals”, *American Economic Review*, vol. 68, nº 2. pp. 364-368.
 165. BUCHANAN, James M.; y TULLOCK, Gordon (1991[1962]), *El cálculo del consenso. Fundamentos lógicos de la democracia constitucional*, capítulo 3, traducido al español por J. Salinas en CABRILLO, F. (comp.), *Lecturas de economía política*, Minerva Ediciones, Madrid. pp. 318-329.
 166. BURLANDO, Roberto; y GUALA, Francesco (2005), “Heterogeneous agents in public goods experiments”, *Experimental Economics*, vol. 8, nº 1. pp. 35-54.
 167. BURNHAM, Terence C. (2003), “Engineering altruism: A theoretical and experimental investigation of anonymity and gift giving”, *Journal of Economic Behavior and Organization*, vol. 50, nº 1. pp. 133-144.
 168. BURNHAM, Terence C.; y JOHNSON, Dominic D. P. (2005), “The biological and evolutionary logic of human cooperation”, *Analyse & Kritik*, vol. 27, nº 1. pp. 113-135.

169. BURNSTEIN, Eugene; CRANDALL, Christian, y KITAYAMA, Shinobu (1994), "Some neo-Darwinian decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision", *Journal of Personality and Social Psychology*, vol. 67, nº 5. pp. 773-789.
170. BUTLER, Joseph (1975[1736]), *The analogy of religion*, Appendix I, en PERRY, John (ed.) (1975), *Personal identity*, University of California Press, Berkeley.
171. BYRNE, Richard W.; y WHITEN, Andrew (1988), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*, Oxford University Press, Oxford.
172. CAHN, B. Rael, DELORME, Arnaud, y POLICH, John (2010), "Occipital gamma activation during Vipassana meditation", *Cognitive Processing*, vol. 11, nº 1. pp. 39-56.
173. CAMERER, Colin F. (2003), *Behavioral game theory. Experiments in strategic interaction*, Princeton University Press, Princeton, New Jersey.
174. CAMERER, Colin F.; y THALER, Richard (1995), "Anomalies: More ultimatum and dictator games", *Journal of Economic Perspectives*, vol. 9, nº 2. pp. 209-219.
175. CAMERER, Colin, F.; LOEWENSTEIN, George; y PRELEC, Drazen (2005), "Neuroeconomics: how neuroscience can inform economics", *Journal of Economic Literature*, vol. 43, nº 1. pp. 9-64.
176. CAMILLERI, Joseph A., KUHLMIEIER, Valerie A., y CHU, Jenny Y. (2010), "Remembering helpers and hinderers depends on behavioral intentions of the agent and psychopathic characteristics of the observer", *Evolutionary Psychology*, vol. 8, nº 2. pp. 303-316.
177. CAMPBELL, Donald T. (1972), "On the genetics of altruism and the counter-hedonic components in human culture", *Journal of Social Issues*, vol. 28, nº 1. pp. 21-37.
178. CAMPBELL, Donald T. (1974), "Evolutionary epistemology", en SCHILPP, Paul A. (ed.), *The philosophy of Karl Popper*, edición en dos volúmenes, Open Court, La Salle, Illinois. pp. 413-463.
179. CAÑIZARES-ESGUERRA, Jorge (2008), *Católicos y puritanos en la colonización de América*, Marcial Pons y Fundación Jorge Juan, Madrid.
180. CAPORAEL, Linnda R.; DAWES, Robyn M.; ORBELL, John; y Van de KRAGT, Alphons J. C. (1989), "Selfishness examined: Cooperation in the absence of egoistic incentives", *Behavioral and Brain Sciences*, vol. 12, nº 4. pp. 683-739.
181. CAPRA, Fritjof (2003), *Las conexiones ocultas. Implicaciones sociales, medioambientales, económicas y biológicas de una nueva visión del mundo*, Editorial Anagrama, Barcelona, 2003.

182. CARLSMITH, Kevin M. (2008), "On justifying punishment: The discrepancy between words and actions", *Social Justice Research*, vol. 21, nº 2. pp. 119-137.
183. CARMICHAEL, M. S.; WARBURTON, V. L.; DIXEN, J.; y DAVIDSON, J. M. (1994), "Relationship among cardiovascular, muscular, and oxytocin responses during human sexual activity", *Archives of Sexual Behavior*, vol. 23, nº 1. pp. 59-79.
184. CARTER, C. Sue (1998), "Neuroendocrine perspectives on social attachment and love", *Psychoneuroendocrinology*, vol. 23, nº 8. pp. 779-818.
185. CASTRO NOGUEIRA, Laureano; LÓPEZ-FANJUL, Carlos; y TORO IBÁÑEZ, Miguel Ángel (2003), *A la sombra de Darwin. Las aproximaciones evolucionistas al comportamiento humano*, prólogo de Francisco Ayala, Editorial Siglo XXI, Madrid.
186. CASTRO NOGUEIRA, Laureano; y TORO IBÁÑEZ, Miguel Ángel (2006), "Evolución y cultura. Los orígenes de la diversidad cultural humana", *Revista de Libros*, nº 118, octubre 2006. Edición en hipertexto: <http://www.revistadelibros.com/articulos/evolucion-y-cultura-los-origenes-de-la-diversidad-cultural-humana>
187. CASTRO, Laureano, y TORO, Miguel A. (2004), "The evolution of culture: From primate social learning to human culture", *Proceedings of the National Academy of Sciences*, vol. 101, nº 27. pp. 10.235-10.240.
188. CAVALIERE, Matteo, SEDWARDS, Sean, TARNITA, Corina E., NOWAK, Martin, A., CSIKÁSZ-NAGY, Attila (2012), "Prosperity is associated with instability in dynamical networks", *Journal of Theoretical Biology*, vol. 299, nº 1. pp. 126-138.
189. CAVALLI-SFORZA, Luigi L. (2010[1996]), *Genes, pueblos y lenguas*, traducción al castellano de Juan Vivanco, Editorial Crítica, Barcelona.
190. CAVALLI-SFORZA, Luigi L.; y FELDMAN, Marcus W.; CHEN, K. H.; y DORNBUSCH, S. M. (1982), "Theory and observation in cultural transmission", *Science*, vol. 218, nº 4.567. pp. 19-27.
191. CHAMPAGNE, Frances A., y JAMES, P. Curley (2005), "How social experiences influence the brain", *Current Opinion in Neurobiology*, vol. 15, nº 6. pp. 704-709.
192. CHAPAIS, Bernard (2010), "The deep structure of human society: Primate origins and evolution", en KAPPELER, Peter M., y SILK, Joan B. (eds.), *Mind the gap. Tracing the origins of human universals*, Springer, Heidelberg. pp. 19-52.
193. CHARNESS, Gary, y RABIN, Matthew (2002), "Understanding social preferences with simple tests", *Quarterly Journal of Economics*, vol. 117, nº 3. pp. 817-869.

194. CHIAPPE, Dan, BROWN, Adam, y DOW, Brian (2004), "Cheaters are looked al longer and remembered better than cooperators in social exchange situations", *Evolutionary Psychology*, vol. 2, n° 1. pp. 108-120.
195. CHIESA, A., y SERRETTI, A. (2010), "A systematic review of neurobiological an clinical features of mindfulness meditations", *Psychological Medicine*, vol. 40, n° 6. pp. 1.239-1.252.
196. CHURCHLAND, Patricia S. (1996), "Feeling reasons", en DAMASIO, A.R., DAMASIO, H., y CHRISTEN, Y. (eds.), *Neurobiology of decision-making*, Springer Verlag, New York. pp. 181-199.
197. CHURCHLAND, Patricia S. (2006), "The big questions: Do we have free wil?", *New Scientist*, n° 2.578. pp. 42-45.
198. CLARK, Kenneth; y SEFTON, Martin (2001), "The secuential prisoner's dilemma: Evidence on reciprocation", *The Economic Journal*, vol. 111, n° 468. pp. 51-68.
199. CLUTIER, Jasmin, KELLEY, William M., y HEATHERTON, Todd (2011), "The influence of perceptual and knowledge-based familiarity on the neural substrates of face perception", *Social Neuroscience*, vol. 6, n° 1. pp. 63-75.
200. CLUTTON-BROCK, T.H., y PARKER, G. A. (1995), "Punishment in animal societies", *Nature*, vol. 373, n° 6.511. pp. 209-216.
201. COHEN Emma (2012), "The evolution of tag-based cooperation in humans: The case for accent", *Current Antropology*, vol. 53, n° 5. pp. 588-616.
202. COHEN-MANSFIELD, J.; GOLANDER, H; y ARNHEIM, G (2000), "Self-identity in older persons suffering from dementia: Preliminary results", *Social Science & Medicine*, vol. 51, n° 3. pp 381-394.
203. COLE, Steve W., HAWKLEY, Louise C., AREVALO, Jesusa M., SUNG, Caroline Y., ROSE, Robert M., CACIOPPO, John T. (2007), "Social regulation of gene expression in human leukocytes", *Genome Biology*, vol. 8, n° 9. R189.
204. COLMAN, Andrew M. (1995), *Game theory and its applications in the social and biological sciences*, 2nd edition, Routledge, London.
205. COLMAN, Andrew M. (2003), "Cooperation, psychological game theory, and limitations of rationality in social interaction", *Behavioral and Brain Sciences*, vol. 26, n° 2. pp. 139-153.
206. COLMAN, Andrew M. (2005), "Evolution of cooperation without awareness in minimal social situations", en GOLD, Nataly (ed.), *Teamwork: Multi-disciplinary perspectives*, Palgrave Macmillan, Basingstoke (UK). pp. 216-235.

207. COLMAN, Andrew M., BROWNING, Lindsay, y PULFORD, Briony D. (2012), "Spontaneous similarity discrimination in the evolution of cooperation", *Journal of Theoretical Biology*, vol. 299, nº 1. pp. 162-171.
208. COLMAN, Andrew M.; PULFORD, Briony D.; y ROSE, Jo (2008a), "Collective rationality in interactive decisions: Evidence for team reasoning", *Acta Psychologica*, vol. 128, nº 2. pp. 387-397. [Target article].
209. COLMAN, Andrew M.; PULFORD, Briony D.; y ROSE, Jo (2008b), "Team reasoning and collective rationality: Piercing the veil of obviousness", *Acta Psychologica*, vol. 128, nº 2. pp. 387-397. [Reply to commentaries].
210. COLMAN, Andrew M.; y BROWNING, Lindsay (2009), "Evolution of cooperative turn-taking", *Evolutionary Ecology Research*, vol. 11, nº 4. pp. 949-963.
211. COOPER, Russell; DeJONG, Douglas V.; FORSYTHE, Robert; y ROSS, Thomas W. (1996), "Cooperation without reputation: Experimental evidence from Prisoner's Dilemma games", *Games and Economic Behavior*, vol. 12, nº 2. pp. 187-218.
212. COPLESTON, Frederick (2011), *Historia de la Filosofía*, edición castellana en cuatro volúmenes dirigida por Manuel Sacristán, Círculo de Lectores, Barcelona.
213. CORICELLI, Giorgio; JOFFILY, Mateus; MONTMARQUETTE, Claude, y VILLEVAL, Marie C. (2010), "Cheating, emotions, and rationality: An experiment on tax evasion", *Experimental Economics*, vol. 13, nº 2. pp. 226-247.
214. COSMIDES, Leda; y TOOBY, John (1992), "Cognitive adaptations for social exchange", en BARKOW, Jerome H; COSMIDES, Leda; y TOOBY, John (eds.), *The adapted mind: Evolutionary psychology and the generation of culture*, Oxford University Press, Oxford & New York. pp. 163-208.
215. COSMIDES, Leda; y TOOBY, John (1994a), "Beyond intuition and instinct blindness: The case for an evolutionary rigorous cognitive science", *Cognition*, vol. 50, nº 1. pp. 41-77.
216. COSMIDES, Leda; y TOOBY, John (1994b), "Better than rational: Evolutionary psychology and the invisible hand", *American Economic Review*, vol. 84, nº 2. pp. 327-332.
217. COSMIDES, Leda; y TOOBY, John (1996), "Are humans good intuitive statisticians after all?: Rethinking some conclusions of the literature of judgment under uncertainty", *Cognition*, vol. 58, nº 1. pp. 1-73.
218. COSMIDES, Leda; y TOOBY, John (2004), "Knowing thyself: The evolutionary psychology of moral reasoning and moral sentiments", en

- FREEMAN, R. Edward, y WERHANE, Patricia H. (eds.), *Business, science, and ethics*, The Ruffin Series of the Society for Business Ethics, vol. 4, Society for Business Ethics, Charlottesville (UK). pp. 93-128.
219. COSMIDES, Leda; y TOOBY, John (2005), “Neurocognitive adaptations designed for social exchange”, en BUSS, David M. (ed.), *The Handbook of Evolutionary Psychology*, Wiley, Hoboken, NJ. pp. 584-627.
220. CRIMMINS, Eileen M., y SEEMAN, Teresa E. (2004), “Integrating biology into the study of health disparities”, en WAITE, Linda J. (ed.), *Aging, health, and public policy: Demographic and economic perspectives*, número suplementario de *Population and Development Review*, vol. 30. pp. 89-107.
221. CUSHMAN, Fiery, y GREENE, Joshua D. (2012), “Finding faults: How moral dilemmas illuminate cognitive structure”, *Social Neuroscience*, vol. 7, nº 3. pp. 269-279.
222. CUYPERS, Stefaan (1993), “Hacia una concepción no atomista de la identidad personal”, *Anuario Filosófico*, vol. 26, pp. 223-248.
223. D'AQUILI, Eugene, y NEWBERG, Andrew B. (1999), *The mystical mind. Probing the biology of religious experience*, Fortress Press, Minneapolis.
224. DALY, Herman E., y COBB, John B. (1994), *For the common good*, 2nd edition, Beacon Press, Boston.
225. DALY, Martin (1978), “The cost of mating”, *The American Naturalist*, vol. 112, nº 986. pp. 771-774. *The American Naturalist*, vol. 30
226. DALY, Martin, y WILSON, Margo (1999), “Human evolutionary psychology and animal behavior”, *Animal Behavior*, vol. 57, nº 4. pp. 509-519.
227. DALY, Martin, y WILSON, Margo (2005), “Human behavior as animal behavior”, en BOLHUIS, J. J., y GIRALDEAU, L. A. (eds.), *Behavior of animals: Mechanisms, function, and evolution*, Blackwell Publishing, Oxford. pp. 393-408.
228. DAMASIO, Antonio R. (2001), *La sensación de lo que ocurre: Cuerpo y emoción en la construcción de la conciencia*, versión castellana de Francisco Páez de la Cadena, Editorial Debate, Madrid.
229. DAMASIO, Antonio R. (2004), *El error de Descartes: La emoción, la razón y el cerebro humano*, traducción castellana de Juandomènec Ros, 3^a edición en castellano, Editorial Crítica, Barcelona.
230. DAMASIO, Antonio R. (2005), *En busca de Spinoza. Neurobiología de la emoción y los sentimientos*, traducción castellana de Juandomènec Ros, Editorial Crítica, Barcelona.

231. DAMASIO, Antonio R., TRANEL, Daniel, y DAMASIO, Hannah (1991), "Somatic markers and the guidance of behavior: Theory and preliminary testing", en LEVIN, H., EISENBERG, H., y BENTON, A. (eds.), *Frontal lobe function and disfunction*, Oxford University Press, New York. pp. 217-229.
232. DARWIN, Charles (1972[1871]), *El origen del hombre, y la selección en relación al sexo*, EDAF, Madrid.
233. DARWIN, Charles (1998a[1859]), *El origen de las especies*, edición abreviada, Editorial Debate, Madrid.
234. DARWIN, Charles (1998b[1872]), *La expresión de las emociones en los animales y en el hombre*, traducción y prólogo de Tomás R. Fernández, Alianza Editorial, Madrid.
235. DASGUPTA, Partha (2000), "Trust as a commodity", en GAMBETTA, Diego (ed.), *Trust: Making and breaking cooperative relations*, Blacwell, Oxford. pp. 49-72.
236. DASGUPTA, Partha (2007), *Economics: A very short introduction*, Oxford University Press, Oxford.
237. DAVIDSON, Donald (1982), "Rational animals", *Dialectica*, vol. 36, nº 4. pp. 317-327.
238. DAVIDSON, Richard J. (2003), "Affective neuroscience and psychophysiology: Toward a Synthesis", *Psychophysiology*, vol. 40, nº 1. pp. 655-665.
239. DAVIS, Daniel (2013), *The compatibility gene*, Allen Lane, London.
240. DAVIS, Douglas D.; y HOLT, Charles A. (1993), *Experimental economics*, Princeton University Press, Princeton.
241. DAVIS, John B. (1995), "Personal identity and standard economic theory", *Journal of Economic Methodology*, vol. 2, nº 1. pp. 35-52.
242. DAVIS, John B. (2006), "Social identity strategies in recent economics", *Journal of Economic Methodology*, vol. 13, nº 3. pp. 371-390.
243. DAVIS, Mark H. (1983), "Measuring individual differences in empathy: Evidence for a multidimensional approach", *Journal of Personality and Social Psychology*, vol. 44, nº 1. pp. 113-126.
244. DAVIS, Mark H. (1994), *Empathy: A social psychological approach*, Brown & Benchmark, Wisconsin.
245. DAVIS, Paul (2004), "Undermining free will", *Foreign Policy*, Special reports on *The world's most dangerous ideas*, nº 144, sep/oct. pp. 36-38
246. DAWES, Robyn M.; y THALER, Richard H. (1988), "Anomalies: Cooperation", *Journal of Economic Perspectives*, vol. 2, nº 3. pp. 187-197.

247. DAWKINS, Richard (1990), *The extended phenotype: The long reach of the gene*, Oxford University Press, Oxford.
248. DAWKINS, Richard (1993[1976]), *El gen egoísta. Las bases biológicas de nuestra conducta*, Salvat Editores, Barcelona.
249. De DREU, Carsten K.W., GREER, Lindred L., Van KLEEF, Gerber A., SHALVI, Shaul, HANDGRAAF, Michael J.J. (2011), “Oxytocin promotes human ethnocentrism”, *Proceedings of the National Academy of Sciences*, vol. 108, nº 4. pp. 1.262-1.266.
250. De QUERVAIN, Dominique J.F., FISCHBACHER, Urs, TREYER, Valerie, SCHELLHAMMER, Melanie, SCHNYDER, Ulrich, BUCH, Alfred, y FEHR, Ernst (2004), “The neural basis of altruistic punishment”, *Science*, vol. 305, nº 5.688. pp. 1.254-1.258.
251. De VISSER, Leonie; HOMBERG, Judith R.; MITSOGIANNIS, Manuela; ZEEB; Fiona D.; RIVALAN, Marion; FITOUSSI, Aurélie; GALHARDO, Vasco; Van den BOS, Ruud; WINSTANLEY, Catherine; y DELLU-HAGEDORN, Françoise (2011), “Rodent versions of the Iowa gambling task: opportunities and challenges for the understanding of decision-making”, *Frontiers in Neuroscience*, vol. 5, article 109. pp. 1-21.
252. De WAAL, Frans B. M. (1997), *Bien natural. Los orígenes del bien y del mal en los humanos y en otros animales*, Editorial Herder, Barcelona.
253. De WAAL, Frans B. M. (2005), “How animals do business”, *Scientific American*, vol. 292, nº 4. pp. 72-79.
254. De WAAL, Frans B. M. (2006), *Primates and philosophers: How morality evolved*, editado por MACEDO, Stephen, y OBER, Josiah; Princeton University Press, Princeton & New York.
255. De WAAL, Frans B. M. (2008), “How selfish an animal? The case of primate cooperation”, en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 63-75.
256. De WAAL, Frans B. M. (2011), *La edad de la empatía*, traducción de Ambrosio García Leal, primera edición en español, Tusquets, Barcelona.
257. De WAAL, Frans B. M.; y BERGER, Michelle L. (2000), “Payment for labour in monkeys”, *Nature*, vol. 404, nº 6.778. p. 563.
258. De WAAL, Frans B. M.; y LUTTRELL, L. M. (1988), “Mechanisms of social reciprocity in three primate species: symmetrical relationships characteristics or cognition?”, *Ethology & Sociobiology*, vol. 9, nº 1. pp. 101-118.
259. DECETY, Jean; y CHAMINADE, Thierry (2003a), “Neural correlates of feeling sympathy”, *Neuropsychologia*, vol. 41, nº 2. pp. 127-138.

260. DECETY, Jean; y CHAMINADE, Thierry (2003b), "When the self represents the other: A new cognitive neuroscience view on psychological identification", *Consciousness & Cognition*, vol. 12, nº 4. pp. 577-596.
261. DECETY, Jean; y JACKSON, Philip (2004), "The functional architecture of human empathy", *Behavioral & Cognitive Neuroscience Reviews*, vol. 3, nº 2. pp. 71-100.
262. DELEUZE, Gilles (1984), *Spinoza: filosofía práctica*, traducción de Antonio Escohotado, Tusquets Editores, Barcelona.
263. DELGADO, M. R., FRANK, R. H., y PHELPS, E. A. (2005), "Perceptions of moral character modulate the neural systems of reward during the trust game", *Nature Neuroscience*, vol. 8, nº 11. pp. 1.611-1.618.
264. DENNETT, Daniel (1971), "Intentional systems", *Journal of Philosophy*, vol. LXVIII, nº 4. pp. 87-106.
265. DENNETT, Daniel C. (1984), "Cognitive wheels: The frame problem of artificial intelligence", en HOOKWAY, Christopher (ed.), *Minds, machines and evolution*, Cambridge University Press, London. pp. 129-151.
266. DENNETT, Daniel C. (1995), *Darwin's dangerous idea*, Penguin Books, London.
267. DENZAU, Arthur T.; y NORTH, Douglass C. (1994), "Shared mental models: Ideologies and institutions", *Kyklos*, vol. 47, nº 1. pp. 3-31.
268. DEUTSCH, David (1999), *La estructura de la realidad*, Editorial Anagrama, Barcelona.
269. DIAMOND, Lisa M. (2001), "Contributions of psychophysiology to research on adult attachment: Review and recommendations", *Personality and Social Psychology Review*, vol. 5, nº 4. pp. 276-295.
270. DICKINSON, Janis L.; y KOENIG, Walter (2003), "Desperately seeking similarity", *Science*, vol. 300, nº 5.627. pp. 1.887-1.890.
271. DIMBERG, U; THUNBERG, M; y ELMEHED, K. (2000), "Unconscious facial reactions to emotional facial expressions", *Psychological Science*, vol. 11, nº 1. pp. 86-89.
272. DOWDING, Keith; y KING, Desmond (1995) (eds.), *Preferences, institutions and rational choice*, Claredon Press, Oxford. pp. 1-19.
273. DREBER, Anna, RAND, David G., FUDENBERG, Drew, y NOWAK, Martin (2008), "Winners don't punish", *Nature*, vol. 452, nº 7.185. pp. 348-351.

274. DRUSKAT, Vanesa U., y WOLFF, Steven B. (2005), "Inteligencia emocional grupal y su influencia en la efectividad del grupo", en GOLEMAN y CHERNISS (eds.) (2005). pp. 203-230.
275. DUGATKIN, Lee A. (2007), *Qué es el altruismo. La búsqueda científica del origen de la generosidad*, Katz Editores, Madrid.
276. DUGATKIN, Lee A., MESTERTON-GIBBONS, Michael, y HOUSTON, Alasdair I. (1992), "Beyond the prisoner's dilemma: Toward models to discriminate among mechanisms of cooperation in nature", *Trends in Ecology and Evolution*, vol. 7, nº 6. pp 202-205.
277. DUGATKIN, Lee Alan, y WILSON, David Sloan (1991), "Rover: A strategy for exploiting cooperators in a patchy environment", *American Naturalist*, vol. 138, nº 3. pp. 678-701.
278. DUNBAR, Robin I. M. (1992), "Neocortical size as a constraint on group size in primates", *Journal of Human Evolution*, vol. 22, nº 6. pp. 469-493.
279. DUNBAR, Robin I. M. (1993), "Coevolution of neocortical size, group size and language in humans", *Behavioral and Brain Sciences*, vol. 16, nº 4. p. 681-735.
280. DUNBAR, Robin I. M. (1996), *Grooming, gossip, and the evolution of language*, Harvard University Press, Cambridge, Massachusetts.
281. DUNN, Elizabeth W., AKNIN, Lara B., y NORTON, Michael (2008), "Spending money on others promotes happiness", *Science*, vol. 319, nº 5.870. pp. 1.687-1.688.
282. DURHAM, William H. (1991), *Coevolution: Genes, culture and human diversity*, Stanford University Press, Stanford.
283. EAKIN, Paul John (2001), "Autobiography, identity, an the fictions of memory", en SCHACTER, Daniel; y SCARRY, Elaine (eds.), *Memory, brain and belief*, Harvard University Press, Cambridge & London. pp. 290-301.
284. EBSTEIN, Richard P., SALOMON, Israel, LERER, Elad, UZEFOVSKY, Florina, SHALEV, Idan, GRITSENKO, Inga, RIEBOLD, Mathias, SALOMON, Shahaf, y YIRMIYA, Nurit (2009), "Arginine, vasopressin, and oxytocin modulate human social behavior", *Annals of the New York Academy of Sciences*, vol. 1.167. pp. 87-102.
285. ECCLES, John (1996), *Evolution of the brain. Creation of the self*, Routledge, London & New York.
286. EDELMAN, Gerard M. (1992), *Bright air, brilliant fire*, Basic Books, New York.
287. EDELMAN, Gerard M. (1995), "Memory and the individual soul: Against silly reductionism", en CORNWALL, John (ed.), *Nature's*

- imagination: The frontiers of scientific vision*, Oxford University Press, Oxford. pp. 200-206.
288. EDELMAN, Gerard M., y TONONI, Giulio (1995), “Neural darwinism: The brain as a selective system”, en CORNWALL, John (ed.), *Nature's imagination: The frontiers of scientific vision*, Oxford University Press, Oxford. pp. 78-100.
289. EDGEWORTH, Francis Y. (2000[1881]), *Psicología matemática*, con estudio preliminar de Manuel Jesús González, traducción de Jordi Pascual, Ediciones Pirámide, Madrid.
290. EFFERSON, Charles; LAVINE, Rafael; y FEHR, Ernst (2008), “The coevolution of cultural groups and ingroup favoritism”, *Science*, vol. 321, nº 5.897. pp. 1.844-1.849.
291. EIBL-EIBESFELDT, Irenäus (1982), “Warfare, Man's Indoctrinability and Group Selection”, *Zeitschrift für Tierpsychologie*, nº 60. pp. 177-198.
292. EIBL-EIBESFELDT, Irenäus (1994), *Amor y odio. Historia natural del comportamiento humano*, Salvat Editores, Barcelona.
293. EIBL-EIBESFELDT, Irenäus (1996), *La sociedad de la desconfianza. Polémica para un futuro mejor*, Herder, Barcelona.
294. EISENBERG, N. (2000), “Empathy and sympathy”, en LEWIS, M.; y HAVILAND-JONES, J. M. (eds.), *Handbook of emotion*, 2nd edition, Guilford Press, New York. pp. 677-691.
295. EKELUND Jr., Robert B.; y TOLLISON, Robert D. (2011), *Economic origins of Roman Christianity*, University of Chicago Press, Chicago & London.
296. EKMAN, Paul (1985), *Telling lies*, Norton, New York.
297. EKMAN, Paul (2003), *Emotions revealed*, Henry Holt & Co., New York.
298. EKMAN, Paul (2005), *Cómo detectar mentiras*, Ediciones Paidós Ibérica, Barcelona.
299. ELLIOTT, John H. (1996[1963]), *La España imperial. 1469-1716*, traducción de J. Marfany, Círculo de Lectores, Barcelona.
300. ELSTER, Jon (1987) (comp.), *The multiple self. Studies in rationality and social change*, Cambridge University Press, Cambridge.
301. ELSTER, Jon (1993), “Why things don't happen as planned”, en AKERMAN, N. (ed.), *The necessity of friction*, Physisca Verlag, Heidelberg. pp. 248-256.
302. ELSTER, Jon (1994), *Justicia local. De qué modo las instituciones distribuyen bienes escasos y cargas necesarias*, traducción de Elena Alterman, primera edición en español, Editorial Gedisa, Barcelona.

303. ELSTER, Jon (1997), *Economics. Análisis de la interacción entre racionalidad, emoción, preferencias y normas sociales en la economía de la acción individual y sus desviaciones*, Editorial Gedisa, Barcelona.
304. ELSTER, Jon (2000[1979]), *Ulises y las sirenas. Estudios sobre racionalidad e irracionalidad*, tercera reimpresión de la primera edición en español, Fondo de Cultura Económica, México.
305. ELSTER, Jon (2001), *Sobre las pasiones. Emoción, adicción y conducta humana*, traducción de J. Francisco Álvarez y Adriana Kiczkowski, Editorial Paidós, Barcelona.
306. EMMONS, Natalie A., y KELEMEN, Deborah (2014), “The developmen of children's prelife reasoning: Evidence from two cultures”, *Child Development*, vol. 85, nº 4. pp. 1.617-1.633.
307. ENGELL, Andrew D., HAXBY, James V., y TODOROV, Alexander (2007), “Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala”, *Journal of Cognitive Neuroscience*, vol. 19, nº 9. pp. 1.508-1.519.
308. ENQUIST, Magnus, y LEIMAR, Olof (1993), “The evolution of cooperation in mobile organisms”, *Animal Behavior*, vol. 45, nº 4. pp. 747-757.
309. EPICTETO (1993), *Disertaciones por Arriano*, traducción, introducción y notas de Paloma Ortíz García, Editorial Gredos, Madrid.
310. ESHEL, Idam, y CAVALLI-SFORZA, Luigi L. (1982), “Assortment of encounters and evolution of cooperativeness”, *Proceedings of the National Academy of Sciences*, vol. 79, nº 3. pp. 687-701.
311. ESPMARK, Yngve, AMUNDSSEN, Trond, y ROSENQVIST, Gunilla (eds.) (2000), *Animal signals: Signaling and signal design in animal communication*, Tapic Academic, Trondheim.
312. ETCOFF, Nancy L., EKMAN, Paul, *et. al.* (2000), “Lie detection and language comprehension”, *Nature*, vol. 405, nº 6.783. pp. 139-141.
313. EVANS, Dylan (2001), *Emotion. The science of sentiment*, Oxford University Press, Oxford & New York.
314. FARRELLI, Daniel, LAZARUS, John, y ROBERTS, Gilbert (2007), “Altruists attract”, *Evolutionary Psychology*, vol. 5, nº 2. pp. 313-329.
315. FEENEY, Brooke, C., y COLLINS, Nancy L. (2001), “Predictions of caregiving in adult intimate relationships: An attachment theoretical perspective”, *Journal of Personality and Social Psychology*, vol. 80, nº 6. pp. 972-994.
316. FEHR, Ernst, GÄCHTER, Simon (2000a), “Cooperation and punishment in public goods experiments”, *American Economic Review*, vol. 90, nº 4. pp. 980-994.

317. FEHR, Ernst, GÄCHTER, Simon (2000b), “Fairness and retaliation: The economics of reciprocity”, *Journal of Economic Perspectives*, vol. 14, nº 3. pp. 159-181.
318. FEHR, Ernst, y CAMERER, Colin F. (2007), “Social neuroeconomics: The neural circuitry of social preferences”, *Trends in Cognitive Sciences*, vol. 11, nº 10. pp. 419-427.
319. FEHR, Ernst, y GÄCHTER, Simon (2002), “Altruistic punishment in humans”, *Nature*, nº 415. pp. 137-140.
320. FEHR, Ernst; FISCHBACHER, Urs, y KOSFELD, Michael (2005), “Neuroeconomic foundations of trust and social preferences: Initial evidence”, *American Economic Review*, vol. 95, nº 2. pp. 346-351.
321. FEHR, Ernst; FISCHBACHER, Urs; y GÄCHTER, Simon (2002), “Strong reciprocity, human cooperation and the enforcement of social norms”, *Human Nature*, vol. 13, nº 1. pp. 1-25.
322. FEHR, Ernst; y FISCHBACHER, URS (2003), “The nature of human altruism”, *Nature*, vol. 425, nº 6.960. pp. 785-791.
323. FEHR, Ernst; y SCHMIDT, Ken (1999), “A theory of fairness, competition and cooperation”, *Quarterly Journal of Economics*, vol. 114, nº 4. pp. 817-868.
324. FEINBERG, Todd E. (2001), *Altered egos: How the brain creates de self*, Oxford University Press, New York.
325. FELDMAN, Marcus W., y CAVALLI-SFORZA, Luigi L. (1989), “On the theory of evolution under genetic and cultural transmission with application to the lactose absorption problem”, en FELDMAN, Marcus W. (ed.), *Mathematical evolutionary theory*, Princeton University Press, Princeton. pp. 145-173
326. FERGUSON, Adam (1974[1767]), *Un ensayo sobre la historia de la sociedad civil*, prólogo de Graciela Soriano, traducción de la cuarta edición de 1773, revisión y corrección de Juan Rincón Jurado, Instituto de Estudios Políticos, Madrid.
327. FERRARO, Fabrizio; PFEFFER, Jeffrey; y SUTTON, Robert I. (2005), “Economics language and assumptions: How theories become self-fulfilling”, *Academy of Management Review*, vol. 30, nº 1. pp. 8-24.
328. FERRÚS, Alberto (1998), “El comportamiento de los genes”, *Mundo Científico*, nº 194. pp. 75-81.
329. FESTINGER, Leon (1957), *A theory of cognitive dissonance*, Stanford University Press, Stanford (CA).
330. FETCHENHAUER, Detlef, GROOTHUIS, Ton, y PRADEL, Julia (2010), “Not only states but traits – Human can identify permanent

- altruistic dispositions in 20 s”, *Evolution and Human Behavior*, vol. 31, n° 2. pp. 80-86.
331. FIDDICK, Laurence, y ERLICH, Nicole (2010), “Giving it all away: Altruism and answers to the Wason selection task”, *Evolution and Human Behavior*, vol. 31, n° 2. pp. 131-140.
332. FISCHBACHER, Urs; GÄCHTER, Simon, y FEHR, Ernst (2001), “Are people conditionally cooperative? Evidence from a public goods experiment”, *Economic Letters*, vol. 71, n° 3. pp. 397-404.
333. FISCHER, Roland (1990), “Why the mind is not in the head but in society's connectionist network?”, *Diogenes*, vol. 35, n° 151. pp. 1-28.
334. FISKE, Alan P. (1991), *Structures of social life. The four elementary forms of human relations*, The Free Press, New York.
335. FISKE, Alan P. (1998), “Human sociality”, *International Society for the Study of Personal Relationships Bulletin*, vol. 14, n° 2. pp. 4-9.
336. FISKE, Alan P. (2002), “Socio-moral emotions motivate action to sustain relationships”, *Self and Identity*, vol. 1, n° 1. pp. 169-175.
337. FISKE, Susan T. (2009), “From dehumanization and objectification to rehumanization. Neuroimaging studies on the building blocks of empathy”, *Annals of the New York Academy of Sciences*, vol. 1.167. pp. 31-34.
338. FLETCHER, Jeffrey A., y DOEBELI, Michael (2009), “A simple and general explanation for the evolution of altruism”, *Proceedings of the Royal Society of London (Biological Sciences)*, vol. 276. pp. 13-19.
339. FONTAINE, Philippe (1997), “Identification and economic behavior. Sympathy and empathy in historical perspective”, *Economics and Philosophy*, vol. 13, n° 2. pp. 261-280.
340. FONTAINE, Philippe (2000), “Making use of the past: Theorists and historians on the economics of altruism”, *European Journal of the History of Economic Thought*, vol. 7, n° 3. pp. 407-422.
341. FONTAINE, Philippe (2007), “From philanthropy to altruism: Incorporating unselfish behavior into economics, 1961-1975”, *History of Political Economy*, vol. 39, n° 1. pp. 1-46.
342. FOOT, Philippa (2002), *Bondad natural. Una vision naturalista de la ética*, Editorial Paidós, Barcelona.
343. FORMIGARI, Lia (1984), *El mono y las estrellas*, traducción de Jorge Larre, Ediciones del Serbal, Barcelona.
344. FOWLER, James H., y CHRISTAKIS, Nicholas A. (2008), “Estimating peer effects on health in social networks”, *Journal of Health Economics*, vol. 27, n° 5. pp. 1.400-1.405.

345. FOWLER, James H., y CHRISTAKIS, Nicholas A. (2010), "Cooperative behavior cascades in human social networks", *Proceedings of the National Academy of Sciences*, vol. 107, nº 12. pp. 5.334-5.338.
346. FOX, Dennis (1985), "Psychology, ideology, utopia, and the commons", *American Psychologist*, vol. 40, nº 1. pp. 48-58.
347. FRANK, Robert H. (1987), "If homo economicus could choose his own utility function, would he want one with a conscience?", *American Economic Review*, vol. 77, nº 4. pp. 593-604.
348. FRANK, Robert H. (1988), *Passions within reason. The strategic role of the emotions*, W. W. Norton & Company, London.
349. FRANK, Robert H. (1994), *Microeconomía y conducta*, traducción de la primera edición en inglés por Luis Toharia y Esther Rabasco; revisión técnica de José Luis Martín Simón, McGraw-Hill, Madrid.
350. FRANK, Robert H. (2001), "Cooperation through emotional commitment", en NESSE, Randolph M. (ed.), *Evolution and the capacity for commitment*, Russell Sage Foundation, New York. pp. 57-76.
351. FRANK, Robert H. (2006), "Cooperation through moral commitment", en BOCK, Greg; y GOODE, Janie (eds.), *Empathy and fairness*, Novartis Foundation Symposium 278, John Wiley & Sons, Chichester (UK). pp. 197-214.
352. FRANK, Robert H. (2008a), *El economista naturalista. En busca de explicaciones para los enigmas cotidianos*, traducción de Carlos Fernández y Victorio Hernández, Ediciones Península, Barcelona.
353. FRANK, Robert H. (2008b), "The status of moral emotions in consequentialist moral reasoning", en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 42-61.
354. FRANK, Robert H.; GILOVICH, Thomas; y REGAN, Dennis T. (1993a), "Does studying economics inhibit cooperation?", *Journal of Economic Perspectives*, vol. 7, nº 2. pp. 159-171.
355. FRANK, Robert H.; GILOVICH, Thomas; y REGAN, Dennis T. (1993b), "The evolution of one-shot cooperation: an experiment", *Ethology and Sociobiology*, vol. 14, nº 2. pp. 247-256.
356. FREDRICKSON, Barbara L., GREWEN, Karen M., COFFEY, Kimberly A., ALGORE, Sara B., FIRESTINE, Ann M., AREVALO, Jesusa M. G., MA, Jeffrey, y COLE, Steven W. (2013), "A functional genomic perspective on human well-being", *Proceedings of the National Academy of Sciences*, vol. 110, nº 33. pp. 13.684-13.689.
357. FREUD, Sigmund (1943[1933(1932)]), "El porqué de la guerra", en *Obras completas*, vol. XVIII, traducción de Ludovico Rosenthal, Editorial Americana, Buenos Aires.

358. FREUD, Sigmund (1997[1929]), *El malestar en la cultura*, con prólogo y álbum de José Rallo, Alianza Editorial, Madrid.
359. FREY, Bruno S. (2006), “Institutions shape preferences: The approach of 'Psychology & Economics'”, en SCHUBERT, Christian, y George von WANGENHEIM (eds.), *Evolution and design of institutions*, Routledge, London. pp. 11-24.
360. FREY, Bruno S., y MEIER, Stephan (2004a), “Pro-social behavior in a natural setting”, *Journal of Economic Behavior and Organization*, vol. 54, nº 1. pp. 65-88.
361. FREY, Bruno S., y MEIER, Stephan (2004b), “Social comparisons and pro-social behavior: Testing 'conditional cooperation' in a field experiment”, *American Economic Review*, vol. 94, nº 5. pp. 1.717-1.722.
362. FREY, Bruno S.; SAVAGE, David A.; y TORGLER, Benno (2010a), “Noblesse oblige? Determinants of survival in a life-and-death situation”, *Journal of Economic Behavior and Organization*, vol. 74, nº 1. pp. 1-11.
363. FREY, Bruno S.; SAVAGE, David A.; y TORGLER, Benno (2010b), “Interaction of natural survival instincts and internalized social norms exploring the Titanic and Lusitania disasters”, *Proceedings of the National Academy of Sciences USA*, vol. 107, nº 11. pp. 4.862-4.865.
364. FREY, Bruno S.; SAVAGE, David A.; y TORGLER, Benno (2010c), “Behavior under extreme conditions: The Titanic disaster”, *Journal of Economic Perspectives*, vol. 24, nº 4. pp. 1-14.
365. FREY, Bruno S.; SAVAGE, David A.; y TORGLER, Benno (2011), “Who perished on the Titanic? The importance of social norms”, *Rationality and Society*, vol. 23, nº 1. pp. 35-49.
366. FRIEDMAN, James W., y HAMMERSTEIN, Peter (1991), “To trade or not to trade, that is the question”, en SELTEN, Reinhard (ed.), *Game equilibrium models*, vol. 1, Springer, Berlin. pp. 257-275.
367. FRIEDMAN, Milton (1967[1953]), “The methodology of positive economics”. Traducción al castellano en *Ensayos sobre economía positiva*, Editorial Gredos, Madrid. pp. 9-44.
368. FRIEDMAN, Milton (1980), *Libertad de elegir. Hacia un nuevo liberalismo económico*, Ediciones Grijalbo, Madrid.
369. FRIEDMAN, Milton (1992), *Paradojas del dinero. Episodios de historia monetaria*, traducción de José Antonio Bravo, Ediciones Grijalbo, Barcelona.
370. FRIEDMAN, Milton; y FRIEDMAN, Rose (1984), *La tiranía del statu quo*, Editorial Ariel, Barcelona.

371. FRITH, Chris D.; y SINGER, Tania (2008), "The role of social cognition in decision making", *Philosophical Transactions of The Royal Society, Biological Sciences*, vol. 363, nº 1.511. pp. 3.785-3.886.
372. FRITH, Uta; y FRITH, Christopher D. (2003), "Development and neurophysiology of mentalizing", *Philosophical Transactions of the Royal Society of London, Biological Sciences*, vol. 358, nº 1.431. pp. 459-473.
373. FUDENBERG, Drew, RAND, David G., y DREBER, Anna (2012), "Slow anger and fast to forgive: Cooperation in an uncertain world", *American Economic Review*, vol. 102, nº 2. Pp. 720-749.
374. FUDENBERG, Drew; y MASKIN, Eric (1986), "The Folk Theorem in repeated games with discounting or with incomplete information", *Econometrica*, vol. 54, nº 3. pp. 533-554.
375. FUKUYAMA, Francis (2000), *La gran ruptura. Naturaleza humana y reconstrucción del orden social*, traducción de Laura Paredes, Ediciones B, Barcelona.
376. FURLONG, Ellen E; y OPFER, John E. (2009), "Cognitive constraints on how economic rewards affect cooperation", *Psychological Science*, vol. 20, nº 1. pp. 11-15.
377. GÄCHTER, Simon; y THÖNI, Christian (2005), "Social learning and voluntary cooperation among like-minded people", *Journal of the European Economic Association*, vol. 3, nº 2-3. pp. 303-314.
378. GALLESE, Vittorio; y GOLDMAN, Alvin (1998), "Mirror neurons and the simulation theory of mind-reading", *Trends in Cognitive Sciences*, vol. 2, nº 12. pp. 493-501.
379. GARCÍA CAMPAYO, Javier (2008), "La práctica de 'estar atento' (*mindfulness*) en medicina. Impacto en pacientes y profesionales", *Atención Primaria*, vol. 40, nº 7. pp. 363-366.
380. GARCÍA LÁZARO, Carlos, CUESTA, José A., SÁNCHEZ, Ángel, y MORENO, Yamir (2012), "Human behavior in prisoner's dilemma experiments suppresses network reciprocity", *Scientific Report*, vol. 2, article number 325. Doi: <http://dx.doi.org/10.1038/srep00325>
381. GARCÍA SUÁREZ, Alfonso (1976), *La lógica de la experiencia. Wittgenstein y el problema del lenguaje privado*, Editorial Tecnos, Madrid.
382. GAT, Azar (2010), "Why war? Motivations for fighting in the human state of nature", en KAPPELER, Peter M., y SILK, Joan B. (eds.), *Mind the gap. Tracing the origins of human universals*, Springer, Heidelberg. pp. 197-220.
383. GAUTHIER, David (1994[1986]), *La moral por acuerdo*, Editorial Gedisa, Barcelona.

384. GAZZANIGA, Michael S. (1993), *El cerebro social*, traducción de Carlos Frade, Alianza Editorial, Madrid.
385. GAZZANIGA, Michael S. (1995), “Principles of human brain organization derived from split-brain studies”, *Neuron*, vol. 14, nº 2. pp. 217-228.
386. GAZZANIGA, Michael S. (1999), *El pasado de la mente*, traducción de Pierre Jacomet, Editorial Andrés Bello, Barcelona.
387. GAZZANIGA, Michael S. (2006), *El cerebro ético*, traducción de Marta Pino Moreno, Ediciones Paidós, Barcelona.
388. GERGEN, Kenneth, J. (1992), *El yo saturado. Dilemas de identidad en el mundo contemporáneo*, Paidós, Barcelona.
389. GHISELIN, Michael T. (1974), *The economy of nature and the evolution of sex*, University of California Press, Berkeley.
390. GIFFORD, Adam (1999), “Being and time: On the nature and the evolution of institutions”, *Journal of Bioeconomics*, vol. 1, nº 2. pp. 127-149.
391. GIGERENZER, Gerd (2000), *Adaptative thinking: Rationality in the real world*, Oxford University Press, New York.
392. GIGERENZER, Gerd (2008), *Decisiones instintivas*, Círculo de Lectores, Barcelona.
393. GIGERENZER, Gerd; TODD, Peter M., y ABC Research Group (eds.) (1999), *Simple heuristics that make us smart*, Oxford University Press, New York.
394. GIL-WHITE, Francisco J. (2001), “Are ethnic groups biological 'species' to the human brain?”, *Current Anthropology*, vol. 42, nº 4. pp. 515-554.
395. GIL-WHITE, Francisco J., y RICHERSON, Peter J. (2003), “Large scale human cooperation and conflict”, en NADEL, Lynn (ed.), *Encyclopedia of Cognitive Science*, Nature Publishing – MacMillan, London. pp. 828-837.
396. GINGES, Jeremy, y ATRAN (2009), “What motivates participation in violent political action: Selective incentives or parochial altruism?”, *Annals of the New York Academy of Sciences*, vol. 1.167, nº 1. pp. 115-123.
397. GINTIS, Herbert (2003), “The hitchhiker’s guide to altruism: Gene-culture coevolution and the internalization of norms”, *Journal of Theoretical Biology*, vol. 220, nº 4. pp. 407-418.
398. GINTIS, Herbert (2006a), “Behavioral ethics meets natural justice”, *Politics, Philosophy & Economics*, vol. 5, nº 1. pp. 5-32.

399. GINTIS, Herbert (2006b), “The flip side of altruism”, entrevista realizada por Chhavi Sachdev en *Altruist International* (3 de abril de 2006).
400. GINTIS, Herbert (2007a), “A framework for the unification of the behavioral sciences” (and Open Peer Commentary) , *Behavioral and Brain Sciences*, vol. 30, nº 1. pp. 1-61.
401. GINTIS, Herbert (2007b), “Adapting minds and evolutionary psychology”, *Journal of Bioeconomics*, vol. 9, nº 2. pp. 191-199.
402. GINTIS, Herbert, HENRICH, Joseph, BOWLES, Samuel, BOYD, Robert, y FEHR, Ernst (2008), “Strong reciprocity and the roots of human morality”, *Social Justice Research*, vol. 21, nº 2. pp. 241-253.
403. GINTIS, Herbert; BOWLES, Samuel; BOYD, Robert; y FEHR, Ernst (2003), “Explaining altruistic behavior in humans”, *Evolution and Human Behavior*, vol. 24, nº 3. pp. 153-172.
404. GINTIS, Herbert; BOWLES, Samuel; BOYD, Robert; y FEHR, Ernst (2005), “Moral sentiments and material interests: Origins, evidence, and consequences”, en GINTIS, Herbert; BOWLES, Samuel; BOYD, Robert; y FEHR, Ernst (eds.), *Moral sentiments and material interests*, MIT Press, Cambridge, MA. pp. 3-39.
405. GINTIS, Herbert; SMITH, Eric A.; y BOWLES, Samuel (2001), “Costly signaling and cooperation”, *Journal of Theoretical Biology*, vol. 213, nº 1. pp. 103-119.
406. GINTIS, Herbert; y KHURANA, Rakesh (2008), “Corporate honesty and business education: A behavioral model”, en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 300-326.
407. GIOCOLI, Nicola (2003), *Modeling rational agents: From interwar economics to early modern game theory*, Edward Elgar, Cheltenham.
408. GLADWELL, Malcolm (2005), *Inteligencia intuitiva*, Taurus, Madrid.
409. GLIMCHER, Paul W. (2003), *Decisions, uncertainty and the brain: The science of neuroeconomics*, MIT Press, Cambridge (Massachusetts).
410. GNEEZY, Ayelet, IMAS, Alex, BROWN, Amber, NELSON, Leif D., y NORTON, Michael (2012), “Paying to be nice: Consistency and costly prosocial behavior”, *Management Science*, vol. 58, nº 1. pp. 179-187.
411. GODE, Dhananjay; y SUNDER, Shyan (1993), “Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality”, *Journal of Political Economy*, vol. 101, nº 1. pp. 119-137.
412. GODFREY-SMITH, Peter (2000), “The replicator in retrospect”, *Biology and Philosophy*, vol. 15, nº 3. pp. 403-423.

413. GOLEMAN, Daniel (2006), *Inteligencia social. La nueva ciencia de las relaciones humanas*, primera edición en castellano, Editorial Kairós, Barcelona.
414. GOOD, David A. (1991), “Cooperation in a microcosm: Lessons from laboratory games”, en HINDE, R. A.; y GROEBEL, J. (eds.), *Cooperation and prosocial behavior*, Cambridge University Press, Cambridge. pp. 224-237.
415. GOODENOUGH, Oliver R. (2008), “Values, mechanism design, and fairness”, en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 228-257.
416. GORDON, Scott (1995), *Historia y filosofía de las ciencias sociales*, traducción de J.M. Álvarez Flórez, Editorial Ariel, Barcelona.
417. GOULD, James M. (1991), “Exadaptation: A crucial tool for an evolutionary psychology”, *Journal of Social Issues*, vol. 47, nº 1. pp. 43-65.
418. GOULD, Stephen Jay (2010[1977]), *Desde Darwin. Reflexiones sobre historia natural*, traducción castellana de Antonio Resines, revisión de Juandomènec Ros, Editorial Crítica, Barcelona.
419. GOULD, Stephen Jay; y VRBA, Elisabeth S. (1982), “Exadaptation – a missing term in the science of form”, *Paleobiology*, vol. 8, nº 1. pp. 4-15.
420. GOULDNER, Alvin (1960), “The norm of reciprocity: A preliminary statement”, *American Sociological Review*, vol. 25, nº 2. pp. 161-178.
421. GREENE, Joshua D. (2003), “From neural 'Is' to moral 'Ought': What are the moral implications of neuroscientific moral psychology?”, *Nature Reviews Neuroscience*, vol. 4, nº 10. pp. 847-850.
422. GREENE, Joshua D. (2004), “Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains”, *TRENDS in Cognitive Sciences*, vol. 11, nº 8. pp. 322-323.
423. GREENE, Joshua D. (2005), “Cognitive neuroscience and the structure of moral mind”, en LAURENCE, S.; CARRUTHERS, P.; y STICH, S. (eds.), *The innate mind: Structure and contents*, Oxford University Press, New York. pp. 338-352.
424. GREENE, Joshua D. (2009), “The cognitive neuroscience of moral judgment”, en GAZZANIGA, Michael S. (ed.), *The cognitive neurosciences*, 4th edition, MIT Press, Cambridge (Massachusetts). pp. 987-1001.
425. GREENE, Joshua D., CUSHMAN, Fiery A., STEWART, Lisa E., LOWENBERG, Kelly, NYSTROM, Leigh E., y COHEN, Jonathan D. (2009), “Pushing moral buttons: The interaction between personal force

- and intention in moral judgement”, *Cognition*, vol. 111, nº 3. pp. 364-371.
426. GREENE, Joshua D., NYSTROM, Leigh E., ENGELL, Andrew D., DARLEY, John M., y COHEN, Jonathan D. (2004), “The neural bases of cognitive conflict and control in moral judgment”, *Neuron*, vol. 44, nº 2. pp. 389-400.
427. GREENE, Joshua D., y BARON, Jonathan (2001), “Intuitions about declining marginal utility”, *Journal of Behavioral Decision Making*, vol. 12, nº 2. pp. 243-255.
428. GREENE, Joshua D., y PAXTON, Joseph M. (2009), “Patterns of neural activity associated with honest and dishonest moral decisions”, *Proceedings of the National Academy of Sciences*, vol. 106, nº 30. pp. 12506-12511.
429. GREENE, Joshua D.; SOMMERVILLE, R. Briam; NYSTROM, Leigh E.; DARLEY, John M.; y COHEN, Jonathan D. (2003), “An fMRI investigation of emotional engagement in moral judgment”, *Science*, vol. 293, nº 5537. pp. 2105-2108.
430. GREGORY, Richard L. (ed.) (2004), “Personal identity”, en *The Oxford Companion to the Mind*, Segunda edición, Oxford University Press, Oxford. pp. 714-716.
431. GREIF, Avner (1994), “Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies”, *Journal of Political Economy*, vol. 102, nº 5. pp. 912-950.
432. GREIF, Avner (2004), “Teoría de juegos e historia de la empresa”, *Información Comercial Española*, nº 812. pp. 9-26.
433. GRUJIC, Jelena, FOSCO, Constanza, ARAUJO, Lourdes, CUESTA, José A., y SÁNCHEZ, Ángel (2010), “Social experiments in the Mesocastle: Humans playing a spatial prisoner's dilemma”, *PloS ONE*, vol. 5, nº 11. e13749.
434. GRUTER, Margaret (1991), *Law and the mind: Biological origins of human behavior*, Sage Publications, Newbury Park (CA).
435. GRZELAK, John (1988), “Conflict and cooperation”, en HEWSTONE, M.; STROEBE, W.; CODOL, J. P.; y STEPHENSON, G. M. (eds.), *Introduction to social psychology: A european perspective*, Basil Blackwell, New York. pp. 288-312.
436. GÜROGLU, Berna, HASELAGER, Gerbert J. T., Van LIESHOUT, Cornelis F. M., TAKASHIMA, Atsuko, RIJKEMA, Mark, y FERNÁNDEZ, Guillén (2008), “Why are friends special? Implementing a social interaction simulation task to probe the neural correlates of friendship”, *Neuroimage*, vol. 39, nº 2. pp. 903-910.

437. GÜTH, Werner; SCHMITTBERGER, Rolf; y SCHWARZE, Bernd (1982), "An experimental analysis of ultimatum bargaining", *Journal of Economic Behavior and Organization*, vol. 3, nº 4. pp. 367-388.
438. GWEON, Hyowon, DODELL-FEDER, David, BEDNY, Marina, y SAXE, Rebecca (2012), "Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts", *Children Development*, vol. 83, nº 6. pp. 1.853-1.868.
439. HALDANE, John B. S. (1932), *The causes of evolution*, Longmans Green, London.
440. HALDANE, John B. S. (1955), "Population genetics", *New Biology*, vol. 18, nº 1. pp. 34-51.
441. HAMANN, Katharina, WARNEKEN, Felix, GREENBERG, Julia R., y TOMASELLO, Michael (2011), "Collaboration encourages equal sharing in children but not in chimpanzees", *Nature*, vol. 476, nº 7.360. pp. 328-331.
442. HAMILTON, William D. (1963), "The evolution of altruistic behaviour", *The American Naturalist*, nº 97. pp. 354-356.
443. HAMILTON, William D. (1964), "The genetical evolution of social behaviour (I-II)", *Journal of Theoretical Biology*, nº 7. pp. 1-52.
444. HAMILTON, William D. (1970), "Selfish and spiteful behaviour in an evolutionary model", *Nature*, vol. 228, nº 277. pp. 1.218-1.220.
445. HAMILTON, William D. (1975), "Innate social aptitudes of man: A approach from evolutionary genetics", en FOX, Robin (ed.), *Biosocial Anthropology*, Wiley, New York. pp. 133-155.
446. HAMILTON, William D., y ZUK, Marlene (1982), "Heritable true fitness and bright birds: A role for parasites?", *Science*, vol. 218, nº 4.570. pp. 384-387.
447. HAMMERSTEIN, Peter (ed.) (2003), *The Genetic and cultural evolution of cooperation*, MIT Press, Cambridge, Massachussets.
448. HARBAUGH, William T.; MAYR, Ulrich; y BURGHART, Daniel R. (2007), "Neural responses to taxation and voluntary giving reveal motives for charitable donations", *Science*, vol. 316, nº 5.831. pp. 1.622-1.625.
449. HARDIN, Garrett (1968), "The tragedy of commons", *Science*, vol. 162, pp. 1.243-1.248.
450. HARDIN, Garrett (1977), *The limits of altruism: An ecologist's view of survival*, Indiana University Press, Bloomington.
451. HARDIN, Russell (2002), *Trust & trustworthiness*, Russell Sage Foundation, New Cork.

452. HARDY-VALLÉE, Benoit (2007), “Decision-making: A neuroeconomic perspective”, *Philosophy Compass*, vol. 2, nº 6. pp. 939-953.
453. HARDY-VALLÉE, Benoit; y POIRIER, Pierre (2006), “Embodied thoughts. Concepts and compositionality without language”, *Theoria et Historia Scientiarum*, vol. 1, nº 1. pp. 53-72.
454. HARE, Robert (1984), *La psicopatía: Teoría e investigación*, Herder, Barcelona.
455. HARE, Robert (2006), *Snakes in suits. When psychopaths go to work*, Regan Books, New York.
456. HAROUSH, Keren, y WILLIAMS, Ziv M. (en prensa), “Neuronal prediction of opponent's behavior during cooperative social interchange in primates”, *Cell* (2015), <http://dx.doi.org/10.1016/j.cell.2015.01.045>
457. HARRISON, Neil A.; SINGER, Tania; ROTSHEIN, Pia; DOLAN, Ray; y CRITCHLEY, Hugo D. (2006), “Pupillary contagion: central mechanisms engaged in sadness processing”, *SCAN*, vol. 1, nº 1. pp. 5-17. doi:10.1093/scan/ns1006
458. HARSANYI, John C. (1955), “Cardinal welfare, individualistic ethics and interpersonal comparisons of utility”, *Journal of Political Economy*, vol. 63, nº 4. pp. 309-321.
459. HAUSER, Marc D. (2006), *Moral minds: How nature designed our universal sense of right and wrong*, Harper Collins, New York.
460. HAUSMAN, Daniel H. (1998), “Rationality and Knavery”, en LEINFELNER, Werner, y KÖHLER, Eckehart (Eds.), *Game theory, experience, rationality. Foundations of social sciences, economics and ethics. In honor of John C. Harsanyi*, Kluwer Academic Publishers, Dordrecht. pp. 67-79.
461. HAWKINS, Jeff (2005), *Sobre la inteligencia*, Espasa-Calpe, Madrid.
462. HAYEK, Friedrich A. (1948), *Individualism and economic order*, University of Chicago Press, Chicago.
463. HAYEK, Friedrich A. (1982[1976]), “Las tres fuentes de la valoración humana”, en *Derecho, legislación y libertad. Una nueva formulación de los principios liberales de la justicia y de la economía política*, vol. 3, traducción de Luis Reig Albiol, Unión Editorial, Madrid. pp. 259-309.
464. HAYEK, Friedrich A. (1990[1988]), *La fatal arrogancia. Los errores del socialismo*, Unión Editorial, Madrid.
465. HAYEK, Friedrich A. (1991[1960]), *Los fundamentos de la libertad*, Unión Editorial, Madrid.
466. HAYEK, Friedrich A. (2003[1952]), *La contrarrevolución de la ciencia. Estudios sobre el abuso de la razón*, Unión Editorial, Madrid.

467. HAYEK, Friedrich A. (2004[1952]), *El orden sensorial: Los fundamentos de la psicología teórica*, con prólogo de Joaquín M. Fuster, Unión Editorial, Madrid.
468. HAYEK, Friedrich A. (2007[1962]), “Reglas, percepción e inteligibilidad”, en *Estudios de filosofía, política y economía*, prólogo de Lorenzo Infantino, traducción de Juan Marcos de la Fuente, Unión Editorial, Madrid. pp. 85-113.
469. HAYEK, Friedrich A. (2007[1967]), “Notas sobre la evolución de los sistemas de reglas de conducta”, en *Estudios de filosofía, política y economía*, prólogo de Lorenzo Infantino, traducción de Juan Marcos de la Fuente, Unión Editorial, Madrid. pp. 115-133.
470. HAYNES, John-Dylan, SAKAI, Katsuyuni, REES, Geraint, GILBERT, Sam, FRITH, Chris, y PASSINGHAM, Richard E. (2007), “Reading hidden intentions in the human brain”, *Current Biology*, vol. 17, nº 4. pp. 323-328.
471. HEBB, Donald (1985[1949]), *Organización de la conducta*, versión española de Tomás del Amo, Editorial Debate, Madrid.
472. HECHTER, Michael; y KANAZAWA, Satoshi (1993), “Group solidarity and social order in Japan”, *Journal of Theoretical Politics*, vol. 5, nº 4. pp. 455-493.
473. HEIN, Grit, y SINGER, Tania (2008), “I feel how you feel but not always: The empathic brain and its modulation”, *Current Opinion in Neurobiology*, vol. 18, nº 2. pp. 153-158.
474. HELLENTHAL, Garrett, BUSBY, George B. J., BAND, Gavin, WILSON, James F., CAPELLI, Cristian, FALUSH, Daniel, y MYERS, Simon (2014), “A genetic atlas of human admixture history”, *Science*, vol. 343, nº 6. pp. 747-751.
475. HELLIWELL, John F.; y BARRINGTON-LEIGH, Christopher (2010), “Measuring and understanding subjective well-being”, National Bureau of Economic Research (NBER) Working Paper Series, nº 15.887, Cambridge (Mass.). Edición en hipertexto: <http://www.nber.org/papers/w15887>
476. HENDERSON, Mark (2010), *50 cosas que hay que saber sobre genética*, traducción de Santiago Madero, Editorial Ariel, Barcelona.
477. HENRICH, J.; BOYD, R.; BOWLES, S.; CAMERER, C.; FEHR, E.; GINTIS, H.; McELREATH, R.; ALVARD, M.; BARR, A.; EUSMINGER, J.; HILL, K.; GIL-WHITE, F.; GURVEN, M.; MARLOWE, F.; PATTON, J. Q.; SMITH, N.; y TRACER, D. (2005), “Economic man in cross-cultural perspective: Behavioral experiments in 15 small-scale societies”, *Behavioral and Brain Sciences*, vol. 28, nº 6. pp. 795-855.

478. HENRICH, Joseph (2004), "Cultural group selection, coevolutionary processes and large-scale cooperation", *Journal of Economic Behavior and Organization*, vol. 53, nº 1. pp. 3-35.
479. HENRICH, Joseph, y GIL-WHITE, Francisco (2001), "The evolution of prestige freely conferred deference as a mechanism for enhancing the benefits of cultural transmission", *Evolution and Social Behavior*, vol. 22, nº 3. pp. 165-196.
480. HENRICH, Joseph; BOYD, Robert; BOWLES, Samuel; CAMERER, Colin; FEHR, Ernst; GINTIS, Herbert; y McELREATH, Richard (2001), "Cooperation, reciprocity and punishment in fifteen small-scale societies", *American Economic Review*, vol. 91, nº 2. pp. 73-78.
481. HENSHILWOOD, C.S., d'ERRICO, F., van NIEKERK, K.A., COAUNOT, Y., JACOBS, Z., LAURITZEN, S.E., MENU, M., y GARCÍA-MORENO, R. (2011), "A 10.000-year-old ochre-processing workshop at Blombos Cave, South Africa", *Science*, vol. 334, nº 6.053. pp. 219-222.
482. HERRANZ, José L. (2013), "El altruismo en las catástrofes", *El País*, 21 de agosto de 2013.
483. HERRMANN-PILLATH, Carsten (1994), "Evolutionary rationality, 'homo economicus', and the foundations of social order", *Journal of Social and Evolutionary Systems*, vol. 117, nº 1. pp. 41-67.
484. HEYES, Cecilia (1998), "Theory of mind in nonhuman primates", *Behavioral and Brain Sciences*, vol. 21, nº 1. pp. 101-148.
485. HIRSHLEIFER, Jack (1978), "Competition, Cooperation, and Conflict in Economics and Biology", *American Economic Review*, vol. 68, nº 2. pp. 238-243.
486. HIRSHLEIFER, Jack (1985), "The expanding domain of economics", *American Economic Review*, vol. 75, nº 6. pp. 53-70.
487. HIRSHLEIFER, Jack (1987), "On the emotions as guarantors of threats and promises", en DUPRÉ, John (ed.), *The latest on the best: Essays on evolution and optimality*, MIT Press, Cambridge, Massachusetts. pp. 307-326.
488. HIRSTEIN, William; y RAMACHANDRAN, Vilallanur S. (1997), "Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons", *Proceedings: Biological Sciences*, vol. 264, nº 1.380. pp. 437-444.
489. HOBBS, Thomas (1993[1651]), *Leviatán*, estudio preliminar y antología de Enrique Tierno Galván, Editorial Tecnos, Madrid.
490. HOFFMAN, Elizabeth; McCABE, Kevin; y SMITH, Vernon L. (1996), "Social distance and other-regarding behavior in dictator games", *American Economic Review*, vol. 86, nº 3. pp. 653-660.

491. HOFFMAN, Elizabeth; McCABE, Kevin; y SMITH, Vernon L. (1998), "Behavioral foundations of reciprocity: Experimental economics and evolutionary psychology", *Economic Inquiry*, vol. 36, n° 3. pp. 335-352.
492. HOFFMAN, Martin L. (2000), *Empathy and moral development. Implications for caring and justice*, Cambridge University Press, Cambridge & New York.
493. HOFSTEDE, Geert (1991), *Cultures and organizations. Software of the mind*, McGraw-Hill, London.
494. HOLROYD, C. B., y COLES, M. G. (2002), "The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity", *Psychological Review*, vol. 109, n° 4. pp. 679-709.
495. HOLT-LUNSTAD, Julianne, SMITH, Timothy B., LAYTON, Bradley J. (2010), "Social relationships and mortality risk: A meta-analytic review", *PloS Medicine*, vol. 7, n° 7, e1000316. pp. 1-20.
496. HOOD, Bruce M., y BLOOM, Paul (2008), "Children prefer certain individuals over perfect duplicates", *Cognition*, vol. 106, n° 1. pp. 455-462.
497. HORIKAWA, T., TAMAKI, M., MIYAWAKI, Y., y KAMITANI, Y. (2013), "Neural decoding of visual imagery during sleep", *Science*, vol. 340, n° 6.132. pp. 639-642.
498. HOWARD, John V. (1988), "Cooperation in the prisoner's dilemma", *Theory and Decision*, vol. 24, n° 3. pp. 203-213.
499. HOWES, David (2003), *Sensual relations: Engaging the senses in culture and social theory*, University of Michigan Press, Ann Arbor (MI).
500. HUCK, Steffen; y TYRAN, Jean-Robert (2007), "Reciprocity, social ties, and competition in markets for experience goods", *Journal of Socio-Economics*, vol. 36, n° 2. pp. 191-203.
501. HULL, David L. (1988), *Science as a process*, Chicago University Press, Chicago.
502. HULL, Glynda A., y ZACHER, Jessica (2007), "Enacting identities: An ethnography of a job training program", *IDENTITY: An International Journal of Theory and Research*, vol. 7, n° 1. pp. 71-102.
503. HUME, David (1992[1739-1740]), *Tratado de la naturaleza humana*, versión española de Vicente Viqueira, Editorial Porrúa, México.
504. HUME, David (1993[1751]), *Investigación sobre los principios de la moral*, traducción, prólogo y notas de Carlos Mellizo, Alianza Editorial, Madrid.
505. HUME, David (1995[1741-42]), "Sobre la dignidad o miseria de la naturaleza humana", en *Sobre el suicidio, y otros ensayos*, selección y

prólogo de Carlos Mellizo, primera reimpresión de la edición española de 1988, Alianza Editorial, Madrid.

506. HURLBERT, Jeanne S.; HAINES, Valerie A.; y BEGGS, John J. (2000), "Core networks and tie activation: What kinds of routine networks allocate resources in non-routine situations", *American Sociological Review*, vol. 65, nº 4. pp. 598-618.
507. HUSTON, Ted L.; y LEVINGER, George (1978), "Interpersonal attraction and relationships", *Annual Review of Psychology*, vol. 29, nº 1. pp. 115-156.
508. HUTH, Alexander G., NISHIMOTO, Shinji, VU, An T., y GALLANT, Jack L. (2012), "A continuous semantic space describes the representation of thousands of object and action categories across the human brain", *Neuron*, vol. 76, nº 6. pp. 1.210-1.224.
509. IACOBONI, Marco (2009), *Las neuronas espejo. Empatía, neuropolítica, autismo, imitación, o de cómo entendemos a los otros*, traducción de Isolda Rodríguez Villegas, Katz Editores, Buenos Aires y Madrid.
510. IACOBONI, Marco, LIEBERMAN, Matthew D., KNOWLTON, Barbara J., MOLNAR-SZAKACS Istvan, MORITZ, Mark, TROOP, C. Jason, y FISKE, Alan P. (2004), "Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline", *NeuroImage*, vol. 21, nº 3. pp. 1.167-1.173.
511. IGLESIAS, María del Carmen (2006), *Razón, sentimiento y utopía*, Círculo de Lectores-Galaxia Gutenberg, Barcelona.
512. IRONS, William (2001), "Religion as a hard-to-fake sign of commitment", en NESSE, Randolph M. (ed.), *Evolution and the capacity for commitment*, Russell Sage Foundation, New York. pp. 292-309.
513. ISAAC, Mark R.; y WALKER, James M. (1988), "Group size effects in public goods provision: The voluntary contributions mechanism", *The Quarterly Journal of Economics*, vol. 103, nº 1. pp. 179-199.
514. JAMES, William (1890), *The principles of psychology*, Macmillan, Londres.
515. JAMES, William (1920[1892]), *Psychology: Briefer course*, Henry Holt & Co., New York.
516. JANSSEN, Marco (2008), "Evolution of cooperation in a one-shot Prisoner's Dilemma based on recognition of trustworthy and untrustworthy agents", *Journal of Economic Behavior and Organization*, vol. 65, nº 3-4. pp. 458-471.
517. JENKINSON, Caroline, DICKENS, Andy P., JONES, Kerry, THOMPSON-COON, Jo, TAYLOR, Rod S., ROGERS, Morwenna,

- BAMBRA, Clare L., LANG, Iain, y RICHARDS, Suzanne H. (2013), "Is volunteering a public health intervention? A systematic review and meta-analysis of the health and survival of volunteers", *BMC Public Health*, vol. 13, nº 773. pp. 1-10.
518. JETTEN, Jolanda; HASLAM, Catherine; PUGLIESE, Cara; TONKS, James; y HASLAM S. Alexander (2010), "Declining autobiographical memory and the loss of identity: Effects on well-being", *Journal of clinical and experimental neuropsychology*, vol. 32, nº 4. pp. 408-416.
519. JEVONS, William Stanley (1998[1871]), *La teoría de la política económica*, estudio preliminar de Manuel Jesús González, Traducción de Juan Pérez-Campanero, revisión de Carlos Rodríguez Braun, Ediciones Pirámide, Madrid.
520. JOHNSON-LAIRD, P. N.; y OATLEY, K. (1992), "Basic emotions, rationality, and folk theory", *Cognition and Emotion*, vol. 6, nº 3-4. pp. 201-223.
521. JONES, Edward; y COLMAN, Andrew M. (1996), "Stereotypes", en KUPER, A; y KUPER, J. (eds.), *The Social Science Encyclopedia*, 2ª edición, Routledge, London. pp. 843-844.
522. JONES, Owen D.; y GOLDSMITH, Timothy H. (2005), "Law and behavioral biology", *Columbia Law Review*, vol. 105, nº 2. pp. 405-502.
523. KAGEL, John H.; BATTALIO, Raymond; RACHIN, Howard; GREEN, Leonard (1981a), "Commodity choice behavior with pigeons as subjects", *Journal of Political Economy*, vol. 89, nº 1. pp. 67-91.
524. KAGEL, John H.; BATTALIO, Raymond; RACHIN, Howard; GREEN, Leonard (1981b), "Demand curves for animal consumers", *Quarterly Journal of Economics*, vol. XCVI, nº 1. pp. 1-15.
525. KAGEL, John H.; BATTALIO, Raymond; RACHIN, Howard; GREEN, Leonard (1995), *Economic choice theory: An experimental analysis of animal behavior*, Cambridge University Press, Cambridge.
526. KAHNEMAN, Daniel; DIENER, Edward; y SCHWARTZ, Norbert (eds.) (1999), *Well-being: The foundations of hedonic psychology*, Russell Sage Foundation, New York.
527. KAHNEMAN, Daniel; GILOVICH, Thomas, y GRIFFIN, Dale (2002) (eds.), *Heuristics and biases: The psychology of intuitive judgment*, Cambridge University Press, Cambridge.
528. KAHNEMAN, Daniel; KNETSCH, Jack L.; y THALER, Richard H. (1991), "The endowment effect, loss aversion, and status quo bias", *Journal of Economic Perspectives*, vol. 5, nº 1. pp. 193-206.
529. KAHNEMAN, Daniel; WALKER, Peter P.; y SARIN, Rakesh (1997), "Back to Bentham? Explorations of experienced utility", *The Quarterly Journal of Economics*, vol. 112, nº 2. pp. 375-405.

530. KANAZAWA, Satoshi (2001), “De gustibus est disputandum”, *Social Forces*, vol. 79, n° 3. pp. 1.131-1.163.
531. KANAZAWA, Satoshi (2004a), “The Savanna Principle”, *Managerial and Decision Economics*, vol. 25, n° 1. pp. 41-54.
532. KANAZAWA, Satoshi (2004b), “General intelligence as a domain-specific adaptation”, *Psychological Review*, vol. 111, n° 2. pp. 512-523.
533. KANAZAWA, Satoshi (2004c), “Social sciences are branches of biology”, *Socio-Economic Review*, vol. 2, n° 3. pp. 371-390.
534. KANDORI, Michihiro (1992), “Social norms and community enforcement”, *Review of Economic Studies*, vol. 59, n° 1. pp. 63-80.
535. KANT, Immanuel (1788[1781]), *Crítica de la razón pura*, con prólogo, traducción, notas e índices de Pedro Ribas, Ediciones Alfaguara, Madrid.
536. KAPLAN, Jonas T., FREEDMAN, Joshua, y IACOBONI, Marco (2007), “Us versus them: Political attitudes and party affiliation influence neural response to faces of presidential candidates”, *Neuropsychologia*, vol. 45, n° 1. pp. 55-64.
537. KELLER, Laurent (ed.) (1999), *Levels of selection in evolution*, Princeton University Press, Princeton.
538. KELLEY, Harold H., y STAHELKY, Anthony J. (1970), “Social interaction basis of cooperators' and competitors' beliefs about others”, *Journal of Personality and Social Psychology*, vol. 16, n° 1. pp. 66-91.
539. KENDLER, Kenneth S., et. al., (2007), “Creating a social world. A developmental twin study of peer-group deviance”, *Archives of General Psychiatry*, vol. 64, n° 8. pp. 958-965.
540. KENNY, Anthony (2005), *Breve historia de la filosofía occidental*, traducción de Miguel Candel, primera edición en castellano, Editorial Paidós, Barcelona.
541. KHALIL, Elias (2004), “What is altruism?”, *Journal of Economic Psychology*, vol. 25, n° 1. pp. 97-123.
542. KIKUCHI, Massao; WATANABE, Yoriko, y YAMAGISHI, Toshio (1996), “Accuracy in the prediction of other's trustworthiness and general trust: an experimental study”, *Japanese Journal of Experimental Social Psychology*, vol. 37, n° 1. pp. 23-36.
543. KILNER, J. M., PAULIGNAN, Y., y BLAKEMORE, S. J. (2003), “An interference effect of observed biological movements on action”, *Current Biology*, vol. 13, n° 6. pp. 522-525.
544. KIMBROUGH, Erik; SMITH, Vernon L.; y WILSON, Bart J. (2008), “Building a market: From personal to impersonal exchange”, en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 280-298.

545. KITCHER, Philip (2006), "Ethics and evolution. How to get here from there", en De WAAL, Frans B. M. (2006), *Primates and philosophers: How morality evolved*, editado por MACEDO, Stephen, y OBER, Josiah; Princeton University Press, Princeton & New York., pp. 125-159.
546. KLEIN, Stanley B.; ROZENDAHL, Keith; y COSMIDES, Leda (2002), "A social-cognitive neuroscience analysis of the self", *Social Cognition*, vol. 20, nº 2. pp. 105-135.
547. KLEIN, Stanley B.; SHERMAN, Jeffrey W.; y LOFTUS, Judith (1996), "The role of episodic and semantic memory in the development of trait self-knowledge", *Social Cognition*, vol. 14, nº 4. pp. 277-291.
548. KNAFO, Ariel, ZAHN-WAXLER, Carolyn, DAVIDOV, Maayan, Van HULLE, Carol, ROBINSON, Joann L., y HYUN RHEE, Soo (2009), "Empathy in early childhood", *Annals of the New York Academy of Sciences*, vol. 1.167. pp. 103-114.
549. KNIGHT, Frank H. (1976[1934]), "Teoría económica del nacionalismo", en *La ética de la sociedad competitiva*, traducción de César Armando Gómez, Unión Editorial, Madrid. pp. 277-377.
550. KOCH, Christoph (2004), *The quest for consciousness. A neurobiological approach*, prólogo de Francis Crick, Roberts and Company Publishers, Englewood.
551. KOENIGS, M.; YOUNG, L.; ADOLPHS, R.; TRANEL, D.; CUSHMAN, F.; HAUSER, M.; DAMASIO, A. (2007), "Damage to the prefrontal cortex increases utilitarian moral judgements", *Nature*, vol. 446, nº 7.138. pp. 908-911.
552. KOLM, Serge-Christophe (1987), "The buddhist theory of 'no-self'", en ELSTER, Jon (1987) (comp.), *The multiple self. Studies in rationality and social change*, Cambridge University Press, Cambridge. pp. 233-265.
553. KOLM, Serge-Christophe (2000a), "Introduction: The economics of reciprocity, giving and altruism", en GÉRARD-VARET, L.-A.; KOLM, S.-C., y MERCIER YTHIER, J. (eds.), *The economics of reciprocity, giving and altruism*, MacMillan Press, London. pp. 1-44.
554. KOLM, Serge-Christophe (2000b), "The theory of reciprocity", en GÉRARD-VARET, L.-A.; KOLM, S.-C., y MERCIER YTHIER, J. (eds.), *The economics of reciprocity, giving and altruism*, MacMillan Press, London. pp. 115-141.
555. KONNER, Melvin (1982), *The tangled wing: Biological constraints on the human spirit*, Harper & Row, New York.
556. KOSFELD, Michael; HEINRICHS, Markus; ZAK, Paul J.; FISCHBACHER, Urs; y FEHR, Ernst (2005), "Oxytocin increases trust in humans", *Nature*, vol. 435, nº 7.042. pp. 673-676.

557. KOSSE, Krisztina (1990), "Group size and societal complexity", *Journal of Anthropological Archaeology*, vol. 9, nº 3. pp. 275-303.
558. KOSTER-HALE, Jorie, SAXE, Rebecca, DUNGAN, James, y YOUNG, Liane L. (2013), "Decoding moral judgements from neural representations of intentions", *Proceedings of the National Academy of Sciences*, vol. 110, nº 14. pp. 5.642-5.647.
559. KRAINES, David; y KRAINES, Vivian (1995), "Evolution of learning among Pavlov strategies in a competitive environment with noise", *Journal of Conflict Resolution*, vol. 39, nº 3. pp. 439-466.
560. KREEK, Mary J.; NIELSEN, David A.; BUTELMAN, Eduardo R., y LaFORGE, K. Steven (2005), "Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction", *Nature Neuroscience*, vol. 8, nº 11. pp. 1.450-1.457.
561. KRETZ, Tobias (2011), "Movimiento de masas", *Mente y Cerebro*, nº 48/2011. pp. 84-89.
562. KROPOTKIN, Piotr A. (1989[1902]), *El apoyo mutuo. Un factor de evolución*, introducción de Ángel J. Cappelletti, Editorial Madre Tierra, Madrid.
563. KROSS, Ethan (2009), "When de self becomes other", *Annals of the New York Academy of Sciences*, vol. 1.167. pp. 35-40.
564. KUNDERA, Milan (2004), *La insoportable levedad del ser*, Tusquets, Barcelona.
565. KUNST-WILSON, William R., y ZAJONC, Robert B. (1980), "Affective discrimination of stimuli that cannot be recognized", *Science*, vol 207, nº 4.430 . pp. 557-558.
566. KURZBAN, Robert, McCABE, Kevin, SMITH, Vernon L., WILSON, Bart J. (2001), "Incremental commitment and reciprocity in a real time public goods game", *Personality and Social Psychology Bulletin*, vol. 27, nº 12. pp. 1.662-1.673.
567. KURZBAN, Robert, y HOUSER, Daniel (2005), "Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations", *Proceedings of the National Academy of Sciences*, vol. 102, nº 5. pp. 1.803-1.807.
568. LAMM, Claus, BATSON, C. Daniel, DECETY, Jean (2007), "The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal", *Journal of Cognitive Neuroscience*, vol. 19, nº 1. pp. 42-58.
569. LANGTON, Cristopher G. (1989), *Artificial Life*, Addison-Wesley, Redwood City, California.

570. LAZARUS, R. S. (1991), "Progress on a cognitive-motivational-relational theory of emotion", *American Psychologist*, vol. 46, nº 8. pp. 819-834.
571. LEAKEY, Richard (2001), *El origen de la humanidad*, traducción de Victoria Laporta, Círculo de Lectores, Barcelona.
572. LeDOUX, Joseph (2000), *El cerebro emocional*, Planeta, Barcelona. Nosotros hemos trabajado la versión inglesa de la obra: *The emotional brain*, Simon & Schuster, New York, 1996.
573. LeDOUX, Joseph; DEBIEC, Jacek; y MOSS, Henry (2003) (dir.), "The self: From soul to brain", número monográfico de *Annals of the New York Academy of Sciences*, vol. 1.001. pp. 1-315.
574. LEDYARD, John O. (1995), "Public goods: A survey of experimental research", en KAGEL, John H.; y ROTH, Alvin E. (eds.), *The handbook of experimental economics*, Princeton University Press, Princeton. pp. 111-194.
575. LEIBERG, Susanne, KLIMECKI, Olga, y SINGER, Tania (2011), "Short-term compassion training increases prosocial behavior in a newly developed prosocial game", *PloS ONE*, vol. 6, nº 3. e17798.
576. LEIDER, Stephen, MÖBIUS, Markus M., ROSENBLAT, Tanya S., y DO, Quoc-Anh (2009), "Directed altruism and enforced reciprocity in social networks", *Quarterly Journal of Economics*, vol. 124, nº 4. pp. 1.185-1.851.
577. LEIDER, Stephen, MÖBIUS, Markus M., ROSENBLAT, Tanya S., y DO, Quoc-Anh (2010), "What do we expect from our friends?", *Journal of the European Economic Association*, vol. 8, nº 1. pp. 120-138.
578. LESTER, Toby (2002), "Oh, Gods!", *Atlantic Monthly*, vol. 289, nº 2. pp. 37-45.
579. LEVI MONTALCINI, Rita (2000), *La galaxia mente*, Editorial Crítica, Barcelona.
580. LÉVI-STRAUSS, Claude (2008[1949]), *Las estructuras elementales del parentesco*, capítulos I – VII, en *Lévi-Strauss. Vida, pensamiento y obra*, extracto de la edición francesa publicada por Mouton & Co. (París – La Haya) traducida al español por Marie Therèse Cevasco, Planeta DeAgostini, Barcelona.
581. LEVINE, David K. (1998), "Modeling altruism and spitefulness in experiments", *Review of Economic Dynamics*, vol. 1, nº 3 pp. 593-622.
582. LEVINE, Sheen S.; y KURZBAN, Robert (2006), "Explaining clustering in social networks: Towards a evolutionary theory of cascading benefits", *Managerial and Decision Economics*, vol. 27, nº 2-3. pp. 173-187.
583. LEVY, Steven (1992), *Artificial life*, Pantheon Books, New York.

584. LEWIS, Michael (2005), "Shared intentions without a self", *Behavioral and Brain Sciences*, vol. 28, n° 5. pp. 707-708.
585. LIBET, Benjamin (1996), "Do we have free will?", *Journal of Consciousness Studies*, vol. 6, n° 8-9. pp. 47-57.
586. LIBET, Benjamin; GLEASON, C.; WRIGHT, E.; y PEARL, D. (1983), "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential)", *Brain*, vol. 106, n° 3. pp. 623-642.
587. LIEBERMAN, Debra; TOOBY, John; y COSMIDES, Leda (2003), "Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest", *Proceedings of the Royal Society of London (Biological Sciences)*, vol. 270. pp. 819-826.
588. LIU, Y., WANG, Z.X. (2003), "Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles", *Neuroscience*, vol. 121, n° 3. pp. 537-544.
589. LLINÁS, Rodolfo (2003), *El cerebro y el mito del yo: El papel de las neuronas en el pensamiento y el comportamiento humano*, traducción de Eugenia Guzmán, Belacqua Ediciones, Barcelona.
590. LOCH, Christoph; GALUNIC, D. Charles; y SCHNEIDER, Susan (2006), "Balancing cooperation and competition in human groups: The role of emotional algorithms and evolution", *Managerial and Decision Economics*, vol. 27, n° 2-3. pp. 217-233.
591. LOCKE, John (1980[1690]), *Ensayo sobre el entendimiento humano*, edición preparada por Sergio Rabade y M^a Esmeralda García, 2 volúmenes, Editora Nacional, Madrid.
592. LOEWENSTEIN, George (1999a), "Experimental economics from the vantage-point of behavioural economics", *The Economic Journal*, vol. 109, n° 453. pp. F25-F34.
593. LOEWENSTEIN, George; HSEE, Christopher K.; WEBER, Elke U., y WELCH, Ned (2001), "Risk as feelings", *Psychological Bulletin*, vol. 127, n° 2. pp. 267-286.
594. LOEWENSTEIN, George; y LERNER, Jennifer S. (2003), "The role of affect in decision making", en DAVIDSON, R. J.; SCHERER, K. R., y GOLDSMITH, H. H., *Handbook of Affective Sciences*, Oxford University Press, Oxford. pp. 619-642.
595. LOEWENSTEIN, George (1996), "Out of control: Visceral influences on behavior", *Organizational Behavior and Human Decision Processes*, vol. 65, n° 3. pp. 272-292.
596. LOMBORG, Bjorn (1996), "Nucleus and shield: Evolution of social structure", *American Sociological Review*, vol. 61, n° 2. pp. 278-307.

597. LÓPEZ CORREDOIRA, Martín (2008), “Naturaleza *versus* libertad”, *Thémata. Revista de Filosofía*, nº 40. pp. 257-266.
598. LÓPEZ MORATALLA, Natalia; y SUEIRO VILLAFRANCA, Enrique (2008), “Células madre y vínculo de apego en el cerebro de la mujer”, Informe Científico sobre la Comunicación Materno-Filial en el Embarazo, Universidad de Navarra. Edición en hipertexto: http://www.unav.es/noticias/natalialopez_cerebroembarazada.pdf
599. LÓPEZ-PÉREZ, Raúl, y SPIEGELMAN, Eli (2013), “Why people tell the truth? Experimental evidence for pure lie aversion”, *Experimental Economics*, vol. 16, nº 3. pp. 233-247.
600. LÓPEZ, Santiago, y VALDALISO, Jesús M^a (1999), “Economía, biología y evolución. Algunas reflexiones sobre la 'Economía evolutiva' y la importancia de la historia”, *Anthropos*, nº 182. pp. 30-41.
601. LORENZ, Konrad (1978[1963]), *Sobre la agresión, el pretendido mal*, octava edición en castellano. Ed. Siglo XXI, México.
602. LORENZ, Konrad (2002[1949]), *El anillo del rey Salomón: Violencia y comunicación en animales y hombres*, traducción de Ramón Margalef, Ediciones Irreverentes y Latorre Literaria, D.L., Madrid.
603. LOTEM, Arnon; FISHMAN, Michael A.; y STONE, Lewi (2003), “From reciprocity to unconditional altruism through signaling benefits”, *Proceedings of the Royal Society of London (Biological Sciences)*, vol. 270. pp. 199-205.
604. LÜSCHER, Christian (2011), “La trampa de la adicción”, *Mente y Cerebro*, nº 48/2011. pp. 26-31.
605. LUTZ, Antoine, SLAGTER, Heleen A., RAWLINGS, Nancy B., FRANCIS, Andrew D., GREISCHAR, Lawrence L., y DAVIDSON, Richard J. (2009), “Mental training enhances attentional stability: Neural and behavioral evidence”, *The Journal of Neuroscience*, vol. 29, nº 42. pp. 13.418-13.427.
606. LUXEN, Marc F., Van de VIJVER, Fons J.R. (2006), “Facial attractiveness, sexual selection, and personnel selection: When evolved preferences matter”, *Journal of Organizational Behavior*, vol. 27, nº 2. pp. 241-225.
607. LYKKEN, David T. (1995), *The antisocial personalities*, Lawrence Erlbaum Associates, Hillsdale (NJ).
608. MA, Ning, VANDEKERCKHOVE, Marie, OVERWALLE, Frank Van; SEURINCK, Ruth, FIAS, Wim (2011), “Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: Spontaneous inferences activate only its core areas”, *Social Neuroscience*, vol. 6, nº 2. pp. 123-138.

609. MACINTYRE, Alasdair (2001a), *Justicia y racionalidad. Conceptos y contextos*, Ediciones Internacionales Universitarias, Madrid.
610. MACINTYRE, Alasdair (2001b), *Animales racionales y dependientes. Por qué los seres humanos necesitamos las virtudes*, Ediciones Paidós, Barcelona.
611. MACINTYRE, Alasdair (2004[1984]), *Tras la virtud*, Ediciones Crítica, Barcelona.
612. MACPHERSON, Crawford B. (1973), *The political theory of possessive individualism: Hobbes to Locke*, tercera reimpression en inglés, Clarendon Press, Oxford.
613. MAGEE, Stephen P. (2006), “Bioeconomics: Lessons for business, nations and life”, en COLANDER, David (ed.), *The complexity vision and the teaching of economics*, Edward Elgar, Cheltenham & Northampton. pp. 255-284.
614. MALINOWSKY, Bronislaw (1926), *Crime and custom in savage society*, International Library of Psychology, Philosophy and Scientific Method, New York.
615. MALONE, Thomas W., y BERNSTEIN, Michael S. (eds.) (2015), *The collective intelligence handbook*, MIT Press, Cambridge, Massachusetts.
616. MARGOLIS, Howard (1982), *Selfishness, altruism and rationality: A theory of social choice*, Cambridge University Press, Cambridge.
617. MARGULIS, Lynn; y FESTER, René (eds.) (1991), *Symbiosis as a source of evolutionary innovation. Speciation and morphogenesis*, MIT Press, Cambridge, Massachusetts.
618. MARTÍN ASUERO, Andrés, GARCÍA de la BANDA, Gloria, BENITO OLIVER, Benito (2005), “Reducción del estrés mediante atención plena: La técnica MBSR en la formación de profesionales de la salud”, *Análisis y Modificación de la Conducta*, vol. 31, nº 139. pp. 557-571.
619. MARTÍN LOZANO, José (1995), “Memoria, quasi-memoria e identidad personal. Una crítica de la teoría de la identidad personal de S. Shoemaker”, *Anales del seminario de metafísica*, nº 29, pp. 73-101.
620. MARTÍN, Victoriano, y SAN EMETERIO, Nieves (2014), “Baltasar Gracián: El concepto de interés propio como guía de la acción humana”, *Estudios de Economía Aplicada*, vol. 32, nº 1. pp. 67-82.
621. MARWELL, Gerard; y AMES, Ruth E. (1981), “Economist free ride: Does anyone else?”, *Journal of Public Economics*, vol. 15, nº 3. pp. 295-310.
622. MARYANSKI, Alexandra, y TURNER, Jonathan H. (1992), *The social cage. Human nature and the evolution of society*, Stanford University Press, Stanford.

623. MASCLET, David; NOUSSAIR, Charles; TUCKER, Steven; y VILLEVAL, Marie-Claire (2003), “Monetary and non-monetary punishment in the voluntary contributions mechanism”, *American Economic Review*, vol. 93, nº 2. pp. 366-380.
624. MASON, Malia F., MAGEE, Joe C., KUWABARA, Ko, y NIND, Louise (2010), “Specialization in relational reasoning: The efficiency, accuracy, and neural substrates of social versus nonsocial inferences”, *Social Psychological and Personality Science*, vol. 1, nº 4. pp. 318-326.
625. MAUSS, Marcel (2002[1923]), *The gift: The form and reason for exchange in archaic societies*, prólogo de Mary Douglas, reedición de la traducción inglesa de Ian Cunnison en 1954 para Routledge, London & New York.
626. MAYNARD SMITH, John (1964), “Group selection and kin selection”, *Nature*, vol. 201, nº 4.924. pp. 1.145-1.147.
627. MAYNARD SMITH, John (1974), “The theory of games and the evolution of animal conflict”, *Journal of Theoretical Biology*, vol. 47, nº 1. pp. 209-221.
628. MAYNARD SMITH, John (1976), “Evolution and the theory of games”, *American Scientist*, vol. 64, nº 1. pp. 41-45.
629. MAYNARD SMITH, John (1982), *Evolution and the theory of games*, Cambridge University Press, Cambridge.
630. MAYNARD SMITH, John; y PRICE, George R. (1973), “The logic of animal conflicts”, *Nature*, vol. 246, nº 5.427. pp. 15-18.
631. MAYNARD SMITH, John; y SZATHMARY, E. (1999), *The origins of life*, Oxford University Press, Oxford.
632. MAYR, Ernst (1961), “Cause and effect in biology”, *Science*, vol. 134, nº 3.489. pp. 1.501-1.506.
633. MAYR, Ernst (2000), *What evolution is*, Basic Books, New York.
634. McCABE, Kevin, HOUSER, Daniel, RYAN, Lee, SMITH, Vernon L., y TROUARD, Theodore (2001), “A functional imaging study of cooperation in two-person reciprocal exchange”, *Proceedings of the National Academy of Sciences*, vol. 98, nº 20. pp. 11.832-11.835.
635. McCABE, Kevin, SMITH, Vernon L., y LePORE, Michael (2000), “Intentionality detection and 'mindreading': Why does game form matter?”, *Proceedings of the National Academy of Sciences*, vol. 97, nº 8. pp. 4.404-4.409.
636. McCLELLAND, David, y KIRCHNIT, Carol (1988), “The effect of motivational arousal through films on salivary immunoglobulin A”, *Psychology and Health*, vol. 2, nº 1. pp. 31-52.

637. McCLINTOCK, Charles G. (1978), "Social motivation -a set of propositions", *Behavioral Science*, vol. 17, n° 5. pp. 438-454.
638. McCULLOUGH, Michael (2008), *Beyond revenge: The evolution of the forgiveness instinct*, John Wiley & Sons, New York.
639. McELREATH, Richard; BOYD, Robert; y RICHERSON, Peter J. (2003), "Shared norms and the evolution of ethnic markers", *Current Anthropology*, vol. 44, n° 1. pp. 122-129.
640. McNEILL, J. R.; y McNEILL, W. H. (2010), *Las redes humanas. Una historia global del mundo*, traducción de Jordi Beltrán, Editorial Crítica, Barcelona.
641. McNULTY, James K., OLSON, Michael A., MELTZER, Andrea L., y SHAFFER, Matthew J. (2013), "Though they may be unaware, newlyweds implicitly know whether their marriage will be satisfying", *Science*, vol. 342, n° 6.162. pp. 1.119-1.120.
642. McPHERSON, Miller; SMITH-LOVIN, Lynn; y COOK, James M. (2001), "Birds of a feather: Homophily in social networks", *Annual Review of Sociology*, vol. 27, n° 1. pp. 415-444.
643. MEALEY, Linda, DAOOD, Christopher, y KRAGE, Michael (1996), "Enhanced memory for faces of cheaters", *Ethology and Sociobiology*, vol. 17, n° 2. pp. 119-128.
644. MEDNICK, Sarnoff A., GABRIELLI, William F., y HUTCHINGS, Barry (1984), "Genetic influence in criminal convictions: Evidence from an adoption cohort", *Science*, vol. 224, n° 4.651. pp. 891-894.
645. MEIER, Stephan (2007), "A survey of economic theories and field evidence on pro-social behavior", en FREY, Bruno S, y STUTZER, Alais (eds.), *Economics and psychology: A promising new cross-disciplinary field*, MIT Press, Cambridge (Mass.) pp. 51-88.
646. MELIS, Alicia P., HARE, Brian, y TOMASELLO, Michael (2006), "Chimpanzees recruit the best collaborators", *Science*, vol. 311, n° 5.765. pp. 1.297-1.300.
647. METZINGER, Thomas (2003), *Being no one: The self-model theory of subjectivity*, MIT Press, Cambridge, Massachusetts.
648. MEYER-LINDENBERG, Andreas, *et. al.* (2006), "Neural mechanisms of genetic risk for impulsivity and violence in humans", *Proceedings of the National Academy of Sciences, USA*, vol. 103, n° 16. pp. 6.269-6.274.
649. MICHOD, Richard E. (1999), *Darwinian dynamics. Evolutionary transitions in fitness and individuality*, Princeton University Press, Princeton.

650. MIFUNE, Nobuhiro, HASHIMOTO, Hirofumi, y YAMAGISHI, Toshio (2010), “Altruism toward in-group members as a reputation mechanism”, *Evolution and Human Behavior*, vol. 31, nº 2. pp. 109-117.
651. MIKULINCER, Mario, y SHAVER, Phillip R. (2001), “Attachment theory and inter-group bias: Evidence that priming the secure base schema attenuates negative reactions to out-groups”, *Journal of Personality and Social Psychology*, vol. 81, nº 1. pp. 97-115.
652. MILGRAM, Stanley (1974), *Obedience to authority: An experimental view*, HarperCollings Publishers, New York.
653. MILGROM, Paul, y ROBERTS, John (1993), *Economía, organización y gestión de la empresa*, traducción de Ernesto Jimeno, primera edición en castellano, Editorial Ariel, Barcelona.
654. MILINSKI, M.; SEMMANN, D.; y KRAMBECK, H. J. (2002), “Reputation helps solve the tragedy of the commons”, *Nature*, vol. 415, nº 6.870. pp. 424-426.
655. MILL, John Stuart (1991[1863]), *El utilitarismo*, introducción, traducción y notas de Esperanza Guisán, Alianza Editorial, Madrid.
656. MILLER, Dale T. (1999), “The norm of self-interest”, *American Psychologist*, vol. 54, nº 12. pp. 1.053-1.060.
657. MILLET, Kobe, y DEWITTE, Siegfried (2007), “Altruistic behavior as a costly signal of general intelligence”, *Journal of Research in Personality*, vol. 41, nº 2. pp. 316-326.
658. MILNER, Richard (1995), *Diccionario de la Evolución. La humanidad a la búsqueda de sus orígenes*, con prólogo de Stephen Jay Gould, traducción de José Luis Gil Aristu, primera edición en español, Biblograf, Barcelona.
659. MISES, Ludwig von (2005[1944]), “El tratamiento de la 'irracionalidad' en las ciencias sociales”, en *Procesos de Mercado: Revista Europea de Economía Política*, vol. II, nº 1, primavera. pp. 199-220. Traducción al español del artículo publicado originalmente en *Philosophy and Phenomenological Research*, vol. 4, nº 4.
660. MISES, Ludwig von (2009[1949]), *La acción humana. Tratado de economía*, novena edición en español, con Estudio preliminar de Jesús Huerta de Soto, traducción de Joaquín Reig Albiol, Unión Editorial, Madrid.
661. MITCHELL, Jason P., BANAJI, Mahzarin R., y MACRAE, C. Neil (2005), “General and specific contributions of the medial prefrontal cortex to knowledge about mental states”, *Neuroimage*, vol. 28, nº 4. pp. 757-762.

662. MITTONE, Luigi, y PLONER, Matteo (2011), "Peer pressure, social spillovers, and reciprocity: An experimental analysis", *Experimental Economics*, vol. 13, nº 2. pp. 203-222.
663. MOLINA, José Luis, FERNÁNDEZ, Rosario, y LLOPIS, Jaime (2008), "El apoyo social en situaciones de crisis: Un estudio de caso desde la perspectiva de las redes personales", *Portularia*, vol. 8, nº 1. pp. 61-76.
664. MOLINA, Vicente (2012), *La psicosis. Ideas sobre la locura*, Editorial Biblioteca Nueva, Madrid.
665. MOLL, Jorge; KRUEGER, Frank; ZAHN, Roland; PARDINI, Matteo; OLIVEIRA-SOUZA, Ricardo de; y GRAFMAN; Jordan (2006), "Human fronto-mesolimbic networks guide decisions about charitable donation", *Proceedings of the National Academy of Sciences*, vol. 103, nº 42. pp. 15.623-15.628.
666. MONTAGE, Read; KING-CASAS, Brooks; y COHEN, Jonathan D. (2006), "Imaging valuation models in human choice", *Annual Review of Neuroscience*, nº 29. pp. 417-448.
667. MONTAGE, Read; y BERNS, Gregory S. (2002), "Neural economics and the biological substrates of valuation", *Neuron*, vol. 36, nº 2. pp. 265-284.
668. MONTAIGNE, Michael de (2010[1580, 1588]), *Ensayos*, edición de María Dolores Picazo, traducción de Almudena Montojo, Cátedra, Madrid.
669. MONTERO, Barbara; y WHITE, Mark (eds.) (2007), *Economics and the mind*, Routledge, Oxon (UK) & New York.
670. MORENO, Juan (2008), *Los retos actuales del darwinismo. ¿Una teoría en crisis*, Editorial Síntesis, Madrid.
671. MORIN, Edgar (2004), *La identidad humana*, traducción de Ana Sánchez, prólogo de Sami Naïr, Círculo de Lectores, Barcelona.
672. MORISHIMA, Yosuke, SCHUNK, Daniel, BRUHIN, Adrian, RUFF, Christian C., y FEHR, Ernst (2012), "Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism", *Neuron*, vol. 75, nº 1. pp. 73-79.
673. MOSTERÍN, Jesús (2006), *La naturaleza humana*, Espasa Calpe, Madrid.
674. MYERSON, Roger B. (1999), "Nash equilibrium and the history of economic theory", *Journal of Economic Literature*, vol. 37, nº 3. pp. 1.067-1.082.
675. NAGEL, Thomas (2004[1970]), *La posibilidad del altruismo*, traducción de Ariel Dilon, primera edición en Español, Fondo de Cultura Económica, México.

676. NANAY, Bence (2010), "Group selection and our obsession with the meaning of life", *The Monist*, vol. 93, nº 1. pp. 76-95.
677. NAQVI, Nasir; SHIV, Baba; y BECHARA, Antoine (2006), "The role of emotion in decision making: A cognitive neuroscience perspective", *Current Directions in Psychological Science*, vol. 15, nº 5. pp. 260-264.
678. NASH, Jonh F. (1950), "The bargaining problem", *Econometrica*, vol. 18, nº 2. pp. 155-162.
679. NASH, Jonh F. (1951), "Non-cooperative games", *Annals of Mathematics*, vol. 54, nº 1. pp. 286-295.
680. NAVARRO, Arcadi (2004), "Conflicto y cooperación en el ámbito humano", ponencia impartida el 19 de julio de 2004 en el taller 3 *Are we selfish or cooperative? Evolutionary roots of social behavior*, incluido en el seminario *The social brain. Biology of conflict and cooperation*, FORUM Barcelona 2004, 17-20 de julio de 2004, Barcelona.
681. NESSE, Randolph M. (2001), "Natural selection and the capacity for subjective commitment", en NESSE, Randolph M. (ed.), *Evolution and the capacity for commitment*, Russell Sage Foundation, New York. pp. 1-44.
682. NEUMANN, John Von, y MORGENSTERN, Oskar (1953[1944]), *The theory of games and economic behavior*, tercera edición, Princeton University Press, Princeton.
683. NICHOLSON, Nigel (1997), "Evolutionary psychology: Toward a new view on human nature and organizational society", *Human Relations*, vol. 50, nº 9. pp. 1.053-1.078.
684. NISBET, Robert (1974), "Cooperación", en SILLS, David L. (dir.), *Enciclopedia Internacional de las Ciencias Sociales*, vol. 3, Editorial Aguilar, Madrid. pp. 182-187.
685. NOË, Ronald; SCHAİK, C.P. van; y HOOFF, Jaram van (1991), "The market effect: An explanation for pay-off asymmetries among collaborating animals", *Ethology*, vol. 87, nº 1. pp. 97-118.
686. NOË, Ronald; y HAMMERSTEIN, Peter (1994), "Biological markets: Supply and demand determine the effect of partner choice in cooperation, mutualism and mating", *Behavioral Ecology and Sociobiology*, vol. 35, nº 1. pp. 1-11.
687. NOE, Thomas N., y REBELLO, Michael J. (1994), "The dynamics of business ethics and economic activity", *American Economic Review*, vol. 84, nº 3. pp. 531-547.
688. NØRRETRANDERS, Tor (1999), *The user illusion. Cutting consciousness down to size*, Penguin Books, New York.

689. NORTH, Douglass C. (1984), *Estructura y cambio en la historia económica*, versión española de M^a Dolores Dionis Trenor y Fernando Fernández Méndez de Andés, Alianza Editorial, Madrid.
690. NORTH, Douglass C. (1993), *Instituciones, cambio institucional y desempeño económico*, primera edición en español de la segunda reimpresión en inglés, Fondo de Cultura Económica, México.
691. NORTH, Douglass C. (2005), *Understanding the process of economic change*, Princeton University Press, Princeton & Oxford.
692. NORTON, Michael I., VANDELLO, Joseph A., y DARLEY, John M. (2004), “Casuistry and social category bias”, *Journal of Personality and Social Psychology*, vol. 87, n° 6. pp. 917-831.
693. NORTON, Michael I., y SOMMERS, Samuel R. (2011), “Whites see racism as a zero-sum game that they are now losing”, *Perspectives on Psychological Science*, vol. 6, n° 3. pp. 215-218.
694. NOWAK, Martin A. (2004), “Prisoners of the dilemma”, *Nature*, vol. 427, n° 6.974. pp. 491-492.
695. NOWAK, Martin A. (2006), “Five rules for the evolution of cooperation”, *Science*, vol. 314, n° 5.805. pp. 1.560-1.563.
696. NOWAK, Martin A., TANITA, Corina E., y WILSON, Edward O. (2010), “The evolution of eusociality”, *Nature*, vol. 466, n° 7.310. pp. 1.057-1.062.
697. NOWAK, Martin A.; MAY, Richard M. (1992), “Evolutionary games and spatial chaos”, *Nature*, vol. 359, n° 6.398. pp. 826-829.
698. NOWAK, Martin A.; PAGE, Karen M.; y SIGMUND, Karl (2000), “Fairness versus reason in the ultimatum game”, *Science*, vol. 289, n° 5.485. pp. 1.773-1.775.
699. NOWAK, Martin A.; SASAKI, Akira; TAYLOR, Christine; y FUDENBERG, Drew (2004), “Emergence of cooperation and evolutionary stability in finite populations”, *Nature*, vol. 428, n° 6.983. pp. 646-650.
700. NOWAK, Martin A.; SIGMUND, Karl (1992), “Tit for tat in heterogeneous populations”, *Nature*, vol. 355, n° 6.357. pp. 250-253.
701. NOWAK, Martin A.; SIGMUND, Karl (1993a), “A strategy of win-stay, lose-shift that outperforms tit for tat in the Prisoner’s Dilemma Game”, *Nature*, vol. 364, n° 6.432. pp. 56-58.
702. NOWAK, Martin A.; SIGMUND, Karl (1993b), “Chaos and the evolution of cooperation”, *Proceedings of the National Academy of Sciences, USA*, vol. 90, n° 11. pp. 5.091-5.094.
703. NOWAK, Martin A.; SIGMUND, Karl (2004), “Evolutionary dynamics of biological games”, *Science*, vol. 303, n° 1.773. pp. 793-799.

704. NOWAK, Martin A.; SIGMUND, Karl (2005), "Evolution of indirect reciprocity", *Nature*, vol. 473, nº 7.063. pp. 1.291-1.298.
705. NOWAK, Martin A.; SIGMUND, Karl; y LEIBOWITZ, Martin L. (2000), "Cooperation versus competition", *Financial Analyst Journal*, vol. 56, nº 4. pp. 13-21.
706. NOWAK, Martin; MAY, Robert M.; y SIGMUND, Karl (1995), "The arithmetics of mutual help", *Scientific American*, nº 6. pp. 76-81.
707. O'BOYLE, Edward, y HARTER, Jim (2013), *State of the global workplace. Employee engagement insights for business leaders worldwide*, Gallup, Washington.
708. OATES, Kerris, y WILSON, Margo (2002), "Nominal kinship cues facilitate altruism", *Proceedings of the Royal Society of London (Biological Sciences)*, vol. 269. pp. 105-109.
709. OBERHOLZER-GEE, Félix (2007), "The helping hand: A brief anatomy", en FREY, Bruno S., y STUTZER, Alais (eds.), *Economics and psychology: A promising new cross-disciplinary field*, MIT Press, Cambridge (Massachusetts). pp. 229-239.
710. OBERMAN, Lindsay M., y RAMACHANDRAN, Vilayanur S. (2007), "The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders", *Psychological Bulletin*, vol. 133, nº 2. pp. 310-327.
711. OBHI, Sukhvinder S., y HAGGARD, Patrick (2005), "Libre albedrío y libre censura", *Mente y cerebro*, nº 11. pp. 15-20.
712. ODA, Ryo (1997), "Biased face recognition in the prisoner's dilemma game", *Evolution and Human Behavior*, vol. 18, nº 5. pp. 309-315.
713. OHTSUKI, Hisashi; y IWASA, Yoh (2004), "How should we define goodness? - Reputation dynamics in indirect reciprocity", *Journal of Theoretical Biology*, vol. 231, nº 1. pp. 107-120.
714. OK, Efe A., y Vega-Redondo, Fernando (2001), "On the evolution of individualistic preferences: An incomplete information scenario", *Journal of Economic Theory*, vo. 97, nº 2. pp. 231-254.
715. OLDS, James; y MILNER, Peter (1954), "Positive reinforcement produced by electrical stimulation of septal area an other regions of rat brain", *Journal of Comparative and Physiological Psychology*, vol. 47, nº 6. pp. 419-427.
716. OLSON, Mancur (1971[1965]), *The logic of collective action: Public goods and theory of groups*, segunda edición, Harvard University Press, Cambridge.
717. OLSON, Mancur (1986), *Auge y decadencia de las naciones*, Ariel, Barcelona.

718. ONES, Umut; y PUTTERMAN, Louis (2007), "The ecology of collective action: A public goods and sanctions experiment with controlled group formation", *Journal of Economic Behavior and Organization*, vol. 62. pp. 495-521.
719. OOSTERBEEK, Hessel; SLOOF, Randolph; y van de KUILEN, Gijs (2004), "Differences in ultimatum game experiments: Evidence from a meta-analysis", *Experimental Economics*, vol. 7, nº 2. pp. 171-188.
720. OPIE, Christopher, ATKINSON, Quentin D., DUNBAR, Robin I. M., y SHULTZ, Susanne (2013), "Male infanticide leads to social monogamy in primates", *Proceedings of the National Academy of Sciences*, vol. 110, nº 33. pp. 13.328-13.332.
721. ORGAN, Dennis W. (1988), *Organizational citizenship behavior. The good soldier syndrome*, D.C. Health & Company, Lexington (Massachusetts).
722. OSTROM, Elinor (1998), "A behavioral approach to the rational choice theory of collective action", *American Political Science Review*, vol. 92, nº 1. pp. 1-22.
723. OSTROM, Elinor (2000), "Collective action and the evolution of social norms", *Journal of Economic Perspectives*, vol. 14, nº 3. pp. 137-158.
724. OSTROM, Elinor (2003), "Toward a behavioral theory linking trust, reciprocity and reputation", en OSTROM, Elinor, y WALKER, James (eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research*, Sage, New York. pp. 19-79.
725. OSTROM, Elinor (2011[1990]), *El gobierno de los bienes comunes. La evolución de las instituciones de acción colectiva*, Fondo de Cultura Económica, México.
726. OSTROM, Elinor, DIETZ, Thomas, DOLSAK, Nives, STERN, Paul C., STONICH, Susan, y WEBER, Elke U. (eds.) (2002), *The drama of the commons*, National Academy Press, Washington, DC.
727. OSTROM, Elinor; WALKER, James; y GARDNER, Roy (1992), "Covenants with and without sword: Self-Governance is possible", *American Political Science Review*, vol. 86, nº 2. pp. 404-417.
728. OYAMA, Susan (2000), *Evolution's eye: A systems view of the biology-culture divide*, Duke University Press, Durham & London.
729. PACHECO, Jorge M.; TRAUlsen, Arne; OHTSUKI, Hisashi; y NOWAK, Martin A. (2008), "Repeated games and direct reciprocity under active linking", *Journal of Theoretical Biology*, vol. 250, nº 4. pp. 723-731.
730. PADGEN, Anthony (2001), *Las tres grandes tradiciones históricas. Una genealogía de la posmodernidad*, Cuaderno nº 8 de la Fundación Juan

March correspondiente al seminario “El pasado y sus críticos”, Fundación Juan March, Madrid.

731. PALFREY, Thomas R., y PRISBREY, Jeffrey E. (1997), “Anomalous behavior in public goods experiments: How much and why?”, *American Economic Review*, vol. 87, nº 4. pp. 829-846.
732. PANCHANATHAN, Karthik; y BOYD, Robert (2004), “Indirect reciprocity can stabilize cooperation without the second-order free rider problem”, *Nature*, vol. 432, nº 7.016. pp. 499-502.
733. PARAMIO, Ludolfo (2000), “Decisión racional y acción colectiva”, *Leviatán*, nº 79. pp. 65-83.
734. PARFIT, Derek (2004[1984]), *Razones y personas*, traducción y estudio introductorio de Mariano Rodríguez González, Antonio Machado Libros, Boadilla del Monte (Madrid).
735. PARFIT, Derek, (1985), “Lewis, Perry y lo que importa”, *Cuadernos de Crítica*, nº 39, versión castellana de Álvaro Rodríguez Tirado, Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, México.
736. PARKER, Geoffrey A. (1984), “Evolutionary stable strategies”, en KREBS, J. R., y DAVIES, N. B. (eds.), *Behavioural ecology: an evolutionary approach*, Blackwell, Oxford. pp. 30-61.
737. PAULSSON, Johan (2002), “Multilevel selection on plasmid replication”, *Genetics*, vol. 161, nº 8. pp. 1.373-1.384.
738. PEDERSEN, Cort A. (2004), “How love evolved from sex and gave birth to intelligence and human nature”, *Journal of Bioeconomics*, vol. 6, nº 1. pp. 39-63.
739. PENNER, Louis A., DOVIDIO, John F., PILIAVIN, Jane A., SCHROEDER, David (2005), “Prosocial behavior: Multilevel perspectives”, *Annual Review of Psychology*, vol. 56. pp. 365-392.
740. PHELPS, Elizabeth A., O'CONNOR, Kevin J., CUNNINGHAM, William A., FUNAYAMA, E. Sumie, GATENBY, J. Christopher, GORE, John, y BANAJI, Mahzarin R. (2000), “Performance on indirect measures of race evaluation predicts amygdala activation”, *Journal of Cognitive Neuroscience*, vol. 12, nº 5. pp. 729-738.
741. PHELPS, Elizabeth, y THOMAS, Laura A. (2003), “Race, behavior, and the brain: The role of neuroimaging in understanding complex social behaviors”, *Political Psychology*, vol. 24, nº 4. pp. 747-758.
742. PIERCE, Barbara; y WHITE, Roderic (1999), “The evolution of social structure: Why biology matters”, *Academy of Management*, vol. 24, nº 4. pp. 843-857.

743. PILIAVIN, Jane A., y CHARNG, Hong-Wen (1990), "Altruism: A review of recent theory and research", *Annual Review of Sociology*, vol. 16. pp. 27-65.
744. PINKER, Steven (1991), *Learnability and cognition: The acquisition of argument structure*, MIT Press, Cambridge.
745. PINKER, Steven (1995), *El instinto del lenguaje*, Alianza Editorial, Madrid.
746. PINKER, Steven (2004), *Cómo funciona la mente*, traducción de Ferran Meler-Otri, segunda edición en español, Ediciones Destino, Barcelona.
747. PINKER, Steven (2005), *La tabla rasa: La negación moderna de la naturaleza humana*, traducción de Roc Filella Escolá, Ediciones Paidós Ibérica, Barcelona.
748. PLACE, Skyler S., TODD, Peter M., PENKE, Lars, y ASENDORPF, Jens B. (2009), "The ability to judge the romantic interest of others", *Psychological Science*, vol. 20, nº 1. pp. 22-26.
749. PLATÓN (1992), *La república o el estado*, vigesimosegunda edición en español, Espasa -Calpe, Madrid.
750. PLOTT, Charles R. (1996), "Rational individual behaviour in markets and social choice processes: The discovered preference hypothesis", en ARROW, Kenneth J.; COLOMBATTO, Enrico; PERLMAN, Mark; y SCHMIDT, Christian (eds.), *The rational foundations of economic behaviour*, Macmillan, London. pp. 225-250.
751. PLUTCHIK, Robert (1980), *Emotion: A psychoevolutionary synthesis*, Harper & Row, New York.
752. POPPER, Karl R. (1991[1979]), *La sociedad abierta y sus enemigos*, traducción de Eduardo Loedel, Ediciones Paidós, Barcelona.
753. POPPER, Karl R. y ECCLES, John (1984), *The self and its brain*, Routledge & Kegan Paul, London.
754. POST, Stephen G. (2009), "It's good to be good: Science says It's so", *Health Progress*, vol. 90, nº 4. pp. 18-25.
755. POUNDSTONE, William (1995), *El dilema del prisionero: John von Neumann, la teoría de juegos y la bomba*, Alianza Editorial, Madrid.
756. PRADO-MARTÍNEZ, Javier, *et. al.* (2013), "Great ape genetic diversity and population history", *Nature*, vol. 499, nº 7.459. pp. 471-475.
757. PRESSMAN, S.D., COHEN, S., MILLER, G.E., BARKIN, A., RABIN, B.S., y TREANOR, J.J. (2005), "Loneliness, social network size, and immune response to influenza vaccination in college freshmen", *Health Psychology*, vo. 24, nº 3. pp. 297-306.

758. PRESTON, Stephanie D.; y De WAAL, Frans B. M. (2002a), "The communication of emotions and the possibility of empathy in animals", en POST, S. G.; UNDERWOOD, L. G.; SCHLOSS, J. P.; y HURLBURT, W. B. (eds.), *Altruistic love: Science, philosophy, and religion in dialogue*, Oxford University Press, Oxford. pp. 284-308.
759. PRESTON, Stephanie D.; y De WAAL, Frans B. M. (2002b), "Empathy: Its ultimate and proximate bases", *Behavioral and Brain Sciences*, vol. 25, nº 1. pp. 1-72.
760. PRICE, Michael E. (2006), "Monitoring, reputation and 'greenbeard' reciprocity in a Shuar work team", *Journal of Organizational Behavior*, vol. 27, nº 2. pp. 201-219.
761. PRICE, Michael E. (2008), "The resurrection of group selection as a theory of human cooperation", *Social Justice Research*, vol. 21, nº 2. pp. 228-240.
762. PRICE, Michael E.; COSMIDES, Leda; y TOOBY, John (2002), "Punitive sentiment as an anti-free rider psychological device", *Evolution and Human Behavior*, vol. 23, nº 2. pp. 203-231.
763. RAIFFA, Howard (1992), "Game theory at the University of Michigan 1948-1952", en WEINTRAUB, E. Roy (ed.), *Toward a history of game theory*, Duke University Press, Durham (North Carolina). Número monográfico de *History of Political Economy*, vol. 24, suplemento anual. pp. 165-175.
764. RAINEY, Paul B.; y RAINEY, Katrina (2003), "Evolution of cooperation and conflict in experimental bacterial populations", *Nature*, vol. 425, nº 6.953. pp. 72-74.
765. RAMACHANDRAN, V. S., y HIRSTEIN, W. (1997), "Three laws of qualia. What neurology tells us about the biological functions of consciousness, qualia and the self", *Journal of Consciousness Studies*, vol. 4, nº 5-6. pp. 429-458.
766. RAMACHANDRAN, Vilayanur S.; y BLAKESLEE, Sandra (1999), *Fantasmas en el cerebro. Los misterios de la mente al descubierto*, versión castellana de Juan Manuel Ibeas, Ediciones Debate, Barcelona.
767. RAMSAY, David (1987), "Hayek's theory of cultural group selection", *Journal of Libertarian Studies*, vol. 8, nº 2. pp. 171-195.
768. RAND, David G., GREENE, Joshua D., NOWAK, Martin (2012), "Spontaneous giving and calculated greed", *Nature*, vol. 489, nº 7.416. pp. 427-430.
769. RAND, David G., y NOWAK, Martin A. (2011), "The evolution of antisocial punishment in optional public goods", *Nature Communications*, vol. 2, nº 434. pp. 1-7.

770. RAND, David, TARNITA, Corina E., OHTSUKI, Hisashi, y NOWAK, Martin (2013), “Evolution of fairness in the one-shot anonymous ultimatum game”, *Proceedings of the National Academy of Sciences*, vol. 110, n° 7. pp. 2.581-2.586.
771. RAPOPORT, Anatol (1989), *Decision theory and decision behavior: Normative and descriptive approaches*, Kluwer Academic Publishers, Dordrecht (Holland).
772. RAPOPORT, Anatol; y CHAMMAH, Albert (1965), *Prisoner's dilemma: A study in conflict and cooperation*, University of Michigan Press, Michigan.
773. RAWLS, John (1972), *A theory of justice*, Oxford University Press, Oxford.
774. REARDON, Sara (2014), “Monkey brains wired to share”, *Nature*, vol. 506, n° 7.489. pp. 416-417.
775. REDCAY, Elizabeth, DODELL-FEDER, David, PEARROW, Mark, MAVROS, Penelope L., y KLEINER, Mario (2010), “Live face to face interaction during fMRI: A new tool for social cognitive neuroscience”, *Neuroimage*, vol. 50, n° 4. pp. 1.639-1.647.
776. RICHARDS, Robert J. (1987), *Darwin and the emergence of evolutionary theories of mind and behavior*, Chicago University Press, Chicago.
777. RICHERSON, Peter J.; y BOYD, Robert (1998), “The evolution of human ultra-sociality”, en EIBL-EIBESFELDT, I.; y SALTER, F. (eds.), *Ideology, warfare and indoctrinability*, Berghan Books, New York. pp. 71-95.
778. RICHERSON, Peter J.; y BOYD, Robert (2001), “The evolution of subjective commitment to groups: A tribal instincts hypothesis” en NESSE, Randolph M. (ed.), *Evolution and the capacity for commitment*, Russell Sage Foundation, New York. pp. 186-220.
779. RICHERSON, Peter J.; y BOYD, Robert (2005), *Not by genes alone: How culture transformed human evolution*, University of Chicago Press, Chicago.
780. RICHERSON, Peter J.; y BOYD, Robert (2008), “The evolution of free enterprise values”, en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 107-141.
781. RIDLEY, Matt (1996), *The origins of virtue. Human instincts and the evolution of cooperation*. Penguin Books, Nueva York.
782. RILLING, James K.; GUTMAN, David A.; ZEH, Thorsten R.; PAGNONI, Guiseppe; BERNS, Gregory S., y KILTS, Clinton D. (2002),

- “A neural basis for social cooperation”, *Neuron*, vol. 35, nº 1. pp. 395-405.
783. RILLING, James K.; SANFEY, Alan G.; ARONSON, Jessica A.; NYSTROM, Leigh E.; y COHEN, Jonathan D. (2004a), “Opposing bold responses to reciprocated and unreciprocated altruism in putative reward pathways”, *Neuroreport*, vol. 15, nº 16. pp. 239-243.
784. RILLING, James K.; SANFEY, Alan G.; ARONSON, Jessica A.; NYSTROM, Leigh E.; y COHEN, Jonathan D. (2004b), “The neural correlates of theory of mind within interpersonal interactions”, *Neuroimage*, vol. 22, nº 4. pp. 1.694-1.703.
785. RIZZOLATTI, Giacomo; y SINIGAGLIA, Corrado (2006), *Las neuronas espejo. Los mecanismos de la empatía emocional*, traducción de Bernardo Moreno Carrillo, Ediciones Paidós Ibérica, Barcelona.
786. ROBBINS, Lionel C. (1934), “Remarks on the relationship between economics and psychology”, *Manchester School of Economics and Social Science*, vol. 5, nº 2. pp. 89-101.
787. ROBBINS, Lionel C. (1935[1932]), *An essay on the nature and significance of economic science*, tercera edición, Macmillan, Londres.
788. ROBERT, Christopher; y CARNEVALE, Peter J. (1997), “Group choice in ultimatum bargaining”, *Organizational Behavior and Human Decision Processes*, vol. 72, nº 2. pp. 256-279.
789. ROBSON, Arthur, J. (2001), “The biological basis of economic behavior”, *Journal of Economic Literature*, vol. 39, nº 1. pp. 11-33.
790. ROBSON, Arthur, J. (2002), “Evolution and human nature”, *Journal of Economic Perspectives*, vol. 16, nº 2. pp. 89-106.
791. ROCKENBACH, Bettina (2007), reseña de GINTIS, Herbert; BOWLES, Samuel; BOYD, Robert; y FEHR, Ernst (2005), *Moral sentiments and material interests*, MIT Press, Cambridge, MA; en *Journal of Economics*, vol. 90, nº 2. pp. 215-218.
792. ROGERS, Alan R. (1994), “Evolution and time preference by natural selection”, *American Economic Review*, vol. 84, nº 3. pp. 460-481.
793. ROSENBERG, Alex; y LINQUIST, Stefan (2005), “On the original contract: Evolutionary game theory and human evolution”, *Analyse & Kritik*, vol. 27, nº 1. pp. 136-157.
794. ROSENFELD, Israel (1986), “Neural darwinism: A new approach to memory and perception”, *The New York Review of Books*, vol. 33, nº 15. Edición en hipertexto: <http://www.nybooks.com/articles/5013>
795. ROSENKRANZ, Melisa A., JACKSON, Daren C., DALTON, Kim M., DOLSKI, Isa, RYFF, Carol D., SINGER, Burt H., MULLER, Daniel, KALIN, Ned H., y DAVIDSON, Richard J. (2003), “Affective style and

- in vivo* immune response: Neurobehavioral mechanisms”, *Proceedings of the National Academy of Sciences*, vol. 100, nº 19. pp. 11.148-11.152.
796. ROSS, Don (2008), “Economics, cognitive science and social cognition”, *Cognitive Systems Research*, vol. 9, nº 1-2. pp. 125-135.
797. ROSSANO, Federico, RAKOCZY, Hannes, y TOMASELLO, Michael (2011), “Young children's understanding of violations of property rights”, *Cognition*, vol. 121, nº 2. pp. 219-227.
798. ROUSSEAU, Jean Jacques (1973[1754]), *Discurso sobre el origen y los fundamentos de la desigualdad entre los hombres*, introducción de Lluís Crespo, traducción de Melitón Bustamante, segunda edición en español, Ediciones Península, Barcelona.
799. RUDOY, John D., y PALLER, Ken A. (2009), “Who can you trust? Behavioral and neural differences between perceptual and memory-based influences”, *Frontiers in Human Neuroscience*, vol. 3, article 16. pp. 1-6. Doi: 10.3389/neuro.09.016.2009
800. RUSE, Michael (1983), *Sociobiología*, Ediciones Cátedra, Madrid.
801. RUSE, Michael (1994), *Tomándose a Darwin en serio. Implicaciones filosóficas del darwinismo*, Salvat Editores, Barcelona.
802. RUSE, Michael (2001), *El misterio de los misterios. ¿Es la evolución una construcción social?*, traducción de Vicente Campos, primera edición en español, Tusquets Editores, Barcelona.
803. RUSSELL, Bertrand (1982[1918]), *Los caminos de la libertad. Socialismo, anarquismo y sindicalismo*, traducción de García Paladini, Ediciones Orbis, Barcelona.
804. RUSTAGI, Devesh; ENGEL, Stefanie, y KOSFELD, Michael (2010), “Conditional cooperation and costly monitoring explain success in forest commons mangement”, *Science*, vol. 330, nº 6.006. pp. 961-965.
805. RYLE, Gilbert (1967[1949]), *El concepto de lo mental*, Paidós, Buenos Aires.
806. SACKS, Oliver (2005), “On memory”, *The Threepenny Review*, vol. 100, winter. Edición en hipertexto: http://threepennyreview.com/samples/sacks_w05.html
807. SAHLINS, Marshall D. (1976), *Uso y abuso de la biología*, traducción de Eulalia Pérez Sedeño, Siglo XXI, Madrid.
808. SAIJO, T.; y NAKAMURA, H. (1995), “The 'spite' dilemma in voluntary contribution mechanism experiments”, *Journal of Conflict Resolution*, vol. 38, nº 3. pp. 535-560.
809. SALLY, David F. (1995), “Conversation and cooperation in social dilemmas. A meta-analysis of experiments from 1958 to 1992”, *Rationality and Society*, vol. 7, nº 1. pp. 58-92.

810. SALLY, David F. (2000), "A general theory of sympathy, mind reading, and social interaction, with an application to the prisoner's dilemma", *Social Science Information*, vol. 29, nº 4. pp. 568-634.
811. SALOM, Julia, y FARINÓS, Joaquín (eds.) (2014), *Identity and territorial character. Re-interpreting local-spatial development*, Universitat de Valencia, Valencia.
812. SAMUELSON, Larry (2005), "Economic theory and experimental economics", *Journal of Economic Literature*, vol. 43, nº 1. pp. 65-107.
813. SAMUELSON, Paul A. (1993), "Altruism as a problem involving group versus individual selection in Economics and Biology", *American Economic Review*, vol. 83, nº 2. pp. 143-155.
814. SANFÉLIX, Vicente (1994), "Las personas y su identidad", *Anales del Seminario de Metafísica*, nº 28, Editorial Complutense, Madrid. pp. 257-285.
815. SANFEY, Alan G.; LOEWENSTEIN, George; McCLURE, Samuel; y COHEN, Jonathan D. (2006), "Neuroeconomics: cross-currents in research on decision-making", *Trends in Cognitive Sciences*, vol. 10, nº 3. pp. 108-116.
816. SANFEY, Alan G.; RILLING, James K.; ARONSON, Jessica A.; NYSTROM, Leigh E.; y COHEN, Jonathan D. (2003), "The neural basis of economic decision-making in the Ultimatum Game", *Science*, vol. 300, nº 5.626. pp. 1.755-1.758.
817. SARTRE, Jean-Paul (2005[1965]), *Bosquejo de una teoría de las emociones*, Alianza Editorial, Madrid.
818. SAUSGRUBER, Rupert (2009), "A note on peer effects between teams", *Experimental Economics*, vol. 12, nº 2. pp. 193-201.
819. SAXE, Rebecca, y CAREY, Susan (2006), "The perception of causality in infancy", *Acta Psychologica*, vol. 123, nº 1-2. pp. 144-165.
820. SAXE, Rebecca, y KANWISHER, N. (2003), "People thinking about thinking people. The role of the temporo-parietal junction in theory of mind", *Neuroimage*, vol. 19, nº 4. pp. 1.835-1.842.
821. SAXE, Rebecca, y WEXLER, Anna (2005), "Making sense of another mind: The role of the right temporo-parietal junction", *Neuropsychologia*, vol. 43, nº 10. pp. 1.391-1.399.
822. SAYGIN, Zeynep M., OSHER, David E., KOLDEWYN, Kami, REYNOLDS, Gretchen, GABRIELI, John D. E., y SAXE, Rebecca (2012), "Anatomical connectivity patterns predict face selectivity in the fusiform gyrus", *Nature Neuroscience*, vol. 15, nº 2. pp. 321-327.

823. SCHACTER, Daniel L. (ed.) (1995), *Memory distortion. How minds, brains and societies reconstruct the past*, Harvard University Press, Cambridge (Mass.).
824. SCHARPF, Friz W. (1990), "Games real actors could play: The problem of mutual predictability", *Rationality and Society*, vol. 2, n° 4. pp. 471-494.
825. SCHEIN, Edgar H. (1987), "Defining organizational culture", en SHAFRITZ, J. M., y OTT, J. S. (eds.), *Classics of organization theory*, The Dorsey Press, Chicago. pp. 373-393.
826. SCHELLING, Thomas C. (1985), "Enforcing rules on oneself", *Journal of Law, Economics, and Organization*, vol. 1, n° 2. pp. 357-374.
827. SCHMID, Hans B. (2003), "Can brains in vats think as a team?", *Philosophical Explorations*, vol. 6, n° 3. pp. 201-218.
828. SCHMID, Hans B. (2005a), "Beyond self-goal choice. Amartya Sen's analysis of the structure of commitment and the role of shared desires", *Economics and Philosophy*, vol. 21, n° 1. pp. 51-63.
829. SCHMID, Hans B. (2005b), "'Nostrism': Social identities in experimental games", *Analyse & Kritik*, vol. 27, n° 1. pp. 172-187.
830. SCHMID, Hans B. (2007), "Towards a strong conception of collective intentionality", en MONTERO, Barbara; y WHITE, Mark (eds.), *Economics and the mind*, Routledge, Oxon (UK) & New York. pp. 159-179.
831. SCHMID, Hans B. (2008), "Plural action", *Philosophy of the Social Sciences*, vol. 38, n° 1. pp. 25-54.
832. SCHNEIDER, Walter E., y SHIFFRIN, Richard M. (1977), "Controlled and automatic human information processing", *Psychological Review*, vol. 84, n° 1. pp. 1-66.
833. SCHOPENHAUER, Arthur (2011[1822]), *Notas sobre Oriente*, edición y estudio de Giovanni Gurisatti, traducción de Adela Muñoz y Paula Caballero, Alianza Editorial, Madrid.
834. SCHUG, Joanna, MATSUMOTO, David, HORITA, Yutaka, YAMAGISHI, Toshio, BONNET, Kemberlee (2010), "Emotional expressivity as a signal of cooperation", *Evolution and Human Behavior*, vol. 31, n° 2. pp. 87-94.
835. SCHULKIN, Jay (2002), "Psychological basis for empathy", *Behavioral and brain Sciences*, vol. 25, n° 1. pp. 46-67.
836. SCHUMPETER, Joseph A. (1939), *Business cycles: A theoretical, historical and statistical analysis of the capitalist process*, 2 volumes, McGraw Hill, New York & London.

837. SCOTT, Alwyn (2000), "How smart is a neuron? A review of Christof Koch's *Biophysics of Computation*", *Journal of Consciousness Studies*, vol. 7, n° 5. pp. 70-75.
838. SEARLE, John R. (2000), *Razones para actuar. Una teoría del libre albedrío*, Círculo de Lectores, Barcelona.
839. SEELEY, Thomas D., VISSCHER, P. Kirk, y PASSINO, Kevin M. (2006), "Group decision making in honey bee swarms", *American Scientist*, vol. 94, n° 3. pp. 220-229.
840. SEFTON, Martin; SHUPP, Robert; y WALKER, James (2002), "The effect of rewards and sanctions in provision of public goods", CeDEX Working Paper 2002-2, University of Nottingham.
841. SEIFRITZ, Enrich, ESPOSITO, Fabrizio, NEUHOFF, John G., LUTHI, Andreas, MUSTOVIC, Hearietta, DAMMAN, Gerhard, Von BARDELEBEN, Ulrich, RADUE, Ernst W., CIRILLO, Sossio, TEDESCHI, Gioacchino, y DISALLE, Francesco (2003), "Sex-independent amigdala response to infant crying and laughing in parents versus nonparents", *Biological Psychiatry*, vol. 54, n° 12. pp. 1.367-1.375.
842. SELL, Jane; y WILSON, Rick K. (1991), "Levels of information and contributions to public goods", *Social Forces*, vol. 70, n° 1. pp. 107-124.
843. SELTEN, Reinhard; y STOECKER, Rolf (1986), "End behavior in sequences of finite Prisoner's Dilemma supergames: A learning theory approach", *Journal of Economic Behavior and Organization*, vol. 7, n° 1. pp. 47-70.
844. SEN, Amartya K. (1977), "Rational fools: A critique of the behavioral foundations of economic theory", *Philosophy and Public Affairs*, vol. 6, n° 4. pp. 317-344.
845. SEN, Amartya K. (1985), "Goals, commitment, and identity", *Journal of Law, Economics and Organization*, vol. 1, n° 2. pp. 341-355.
846. SEN, Amartya K. (1986), "Rationality, interest and identity", en FOXLEY, A.; McPHERSON, M.; y O'DONELL G. (eds.), *Development, democracy and the art of trespassing*, University of Notre Dame Press, Notre Dame. pp. 343-353.
847. SEN, Amartya K. (1989), "A conversation with Amartya Sen", *Journal of Economic Perspectives*, vol. 3, n° 1. pp. 135-150.
848. SEN, Amartya K. (2007), *Identidad y violencia. La ilusión del destino*, traducción de Verónica Inés Weinstabl y Servanda María de Hagen, Katz Editores, Buenos Aires (Argentina).
849. SETHI, Rajiv, y SOMANATHAN, E. (2003), "Understanding reciprocity", *Journal of Economic Behavior and Organization*, vol. 50, n° 1. pp. 1-28.

850. SEYMOUR, Ben; DAW, Nathaniel; DAYAN, Peter; SINGER, Tania; y DOLAN, Ray (2007), "Differential encoding of losses and gains in the human striatum", *Journal of Neuroscience*, vol. 27, n° 18. pp. 4.826-4.831.
851. SHELDON, Kennon M., y HOUSER-MARKO, Linda (2001), "Self-concordance, goal attainment, and pursuit of happiness: Can there be an upward spiral?", *Journal of Personality and Social Psychology*, vol. 80, n° 1. pp. 152-165.
852. SHENKER, Barry (1986), *Intentional communities: Ideology and alienation in communal societies*, Routledge & Kegan Paul, London.
853. SHIV, Baba; LOEWENSTEIN, George; BECHARA, Antoine; DAMASIO, Hanna, y DAMASIO, Antonio R. (2005), "Investment behavior and the dark side of emotion", *Psychological Science*, vol. 16, n° 6. pp. 435-439.
854. SHIZGAL, Peter (2012), "Scarce means with alternative uses: Robbin's definition of economics and its extension to the behavioral and neurobiological study of animal decision making", *Frontiers in Decision Neuroscience*, vol. 6, n° 20. doi: 10.3389/fnins.2012.00020
855. SIDOWSKY, J. B. (1957), "Reward and punishment in the Minimal Social Situation", *Journal of Experimental Psychology*, vol. 54, n° 3. pp. 318-326.
856. SIGMUND, Karl (1993), *Games of life: Explorations in ecology, evolution, and behavior*, Oxford University Press, Oxford.
857. SIGMUND, Karl; FEHR, Ernst; y NOWAK, Martin A. (2002), "The economics of fair play", *Scientific American*, vol. 286, n° 1. pp. 83-87.
858. SIGMUND, Karl; y NOWAK, Martin A. (1997), "The natural history of mutual aid: An eye for an eye, and a meal for a meal", en STADLER, F. (ed.), *Wissenschaft als kultur*, Spinger Verlag, New York. pp. 259-272.
859. SIGMUND, Karl; y NOWAK, Martin A. (2000), "A tale of two selves", *Science*, vol. 290, n° 5.493. pp. 949-950.
860. SILK, Joan B. (2007), "Social components of fitness in primate groups", *Science*, vol. 317, n° 5.843. pp. 1.347-1.351.
861. SILK, Joan B., y BOYD, Robert (2010), "From grooming to giving blood: The origins of human altruism", en KAPPELER, Peter M., y SILK, Joan B. (eds.), *Mind the gap. Tracing the origins of human universals*, Springer, Heidelberg. pp. 223-244.
862. SIMON, Herbert A. (1978), "Rationality as a process and as a product of thought", *American Economic Review*, vol. 68, n° 2. pp. 1-15.
863. SIMON, Herbert A. (1983), *Reason in human affairs*, Stanford University Press, Stanford.

864. SIMON, Herbert A. (1989), *Naturaleza y límites de la razón humana*, Fondo de Cultura Económica, México.
865. SIMON, Herbert A. (1990), “A mechanism for social selection and successful altruism”, *Science*, vol. 250, n° 4.988. pp. 1.665-1.668.
866. SIMON, Herbert A. (1993), “Altruism and economics”, *American Economic Review*, vol. 83, n° 2. pp. 156-161.
867. SIMON, Herbert A. (2001), “Complex systems: The interplay of organizations and markets in contemporary society”, *Computational & Mathematical Organization Theory*, vol. 7, n° 2. pp. 79-85.
868. SINERVO, Barry; y CLOBERT, Jean (2003), “Morphs, dispersal behavior, genetic similarity, and the evolution of cooperation”, *Science*, vol. 300, n° 5.627. pp. 1.949-1.952.
869. SINGER, Peter (2000), *Una izquierda darwiniana. Política, evolución y cooperación*, traducción de A. J. Desmots, Editorial Crítica, Barcelona.
870. SINGER, Tania (2006), “The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research”, *Neuroscience and Biobehavioral Reviews*, vol. 30, n° 6. pp. 855-863.
871. SINGER, Tania, BIRD, Geoffrey, PETROVIC, Predrag, SILANI, Giorgina, y HEINRICHS, Markus (2008), “Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain”, *Emotion*, vol. 8, n° 6. pp. 781-791.
872. SINGER, Tania, SEYMOUR, Ben, O'DOHERTY, John, KAUBE, Holger, DOLAN, Raymond J., y FRITH, Chris (2004a), “Empathy for pain involves the affective but not sensory components of pain”, *Science*, vol. 303, n° 5.661. pp. 1.157-1.162.
873. SINGER, Tania; KIEBEL, S. J.; WINSTON, J. S.; KAUBE, H.; DOLAN, R. J.; y FRITH, C. D. (2004b), “Brain responses to the acquired moral status of faces”, *Neuron*, vol. 41, n° 4. pp. 653-662.
874. SINGER, Tania; SEYMOUR, Ben; O'DOHERTY, John; STEPHAN, Klass E.; DOLAN, Raymond J.; y FRITH, Chris D. (2006), “Empathic neural responses are modulated by the perceived fairness of others”, *Nature*, vol. 439, n° 7.075. pp. 466-469.
875. SINGER, Tania; y KRAFT, Ulrich (2005), “Empatía”, *Mente y Cerebro*, n° 11/2005. pp. 60-65.
876. SINGER, Tania; y LAMM, Claus (2009), “The social neuroscience of empathy”, *Annals of the New York Academy of Sciences*, vol. 1.156. pp. 81-96.
877. SINGER, Tania; y STEINBEIS, Nikolaus (2009), “Differential roles of fairness -and compassion- based motivations for cooperation, defection

- and punishment”, *Annals of the New York Academy of Sciences*, vol. 1.167. pp. 41-50.
878. SKUSE, David H., LORI, Adriana, CUBELLS, Joseph F., LEE, Irene, CONNEELY, Karen N., PUURA, Kaija, LEHTIMÄKI, Terho, BINDER, Elisabeth B., y YOUNG, Larry J. (2014), “Common polymorphism in the oxytocin receptor gene (OXTR) is associated with human social recognition skills”, *Proceedings of the National Academy of Sciences*, vol. 111, nº 5. pp. 1.987-1.992.
879. SMITH, Adam (1978[1762]), *Lectures in Jurisprudence*, Oxford University Press, Oxford.
880. SMITH, Adam (1988[1776]), *Investigación sobre la naturaleza y las causas de la riqueza de las naciones*, Oikos-Tau, Barcelona.
881. SMITH, Adam (1997[1759]), *La teoría de los sentimientos morales*, edición española con estudio preliminar de Carlos Rodríguez Braun, Alianza Editorial, Madrid.
882. SMITH, Adam (1998[1795]), *Ensayos filosóficos. Con una Relación de la vida y escritos de Adam Smith, por Dugald Stewart*, estudio preliminar de John Reeder y traducción de Carlos Rodríguez Braun, Ediciones Pirámide, Madrid.
883. SMITH, Kerry (2013), “Reading minds”, *Nature*, vol. 502, nº 7.472. pp. 428-430.
884. SMITH, Vernon L. (1998), “The two faces of Adam Smith”, *Southern Economic Journal*, vol. 65, nº 1. pp. 2-19.
885. SMITH, Vernon L. (2002), “Constructivist and ecological rationality in economics”, Nobel Prize Lecture, december 8th, 2002. pp. 502-561. Edición en hipertexto:
http://nobelprize.org/nobel_prizes/economics/laureates/2002/smith-lecture.pdf
886. SNYDER, Mark, STUKAS, Arthur A. (1999), “Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction”, *Annual Review of Psychology*, vol. 50, nº 1. pp. 273-303.
887. SOBEL, Joel (2005), “Interdependent preferences and reciprocity”, *Journal of Economic Literature*, vol. 43, nº 2. pp. 392-436.
888. SOBER, Elliott (1996), *Filosofía de la biología*, versión española de Tomás R. Fernández y Susana del Viso, Alianza Editorial, Madrid.
889. SOBER, Elliott, y WILSON, David Sloan (2000), *El comportamiento altruista. Evolución y psicología*, prólogo de Carlos López Fanjul, Siglo Veintiuno de España Editores, Madrid.

890. SOLOMON, Robert (2008), "Free enterprise, sympathy and virtue", en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in economy*, Princeton University Press, Princeton. pp. 16-40.
891. SOMMER, Monika, ROTHMAYR, Christoph, DÖHNEL, Katrin, MEINHARDT, Jörg, SCHWERDTNER, Johannes, SODIAN, Beate, y HAJAK, Göran (2010), "How should I decide? The neural correlates of everyday moral reasoning", *Neuropsychologia*, vo. 48, nº 7. pp. 2.018-2.026.
892. SOSIS, Richard (2005), "Una interpretación darwinista del fenómeno religioso", *Mente y Cerebro*, nº 12/2005. pp. 72-76.
893. SPERBER, Dan (2001), "Lo individual bajo el peso de lo colectivo", *Mundo Científico*, nº 227. pp. 28-31.
894. SPINOZA, Baruch (1988[1674]), "Carta a G. H. Schuller" (carta nº 58), en *Spinoza. Correspondencia*, introducción, traducción, notas e índices de Atilano Domínguez, Alianza Editorial, Madrid. pp. 265-266.
895. SPINOZA, Baruch (2004[1677]), *Ética demostrada según el orden geométrico*, introducción, traducción y notas de Vidal Peña, Alianza Editorial, Madrid.
896. SRIVASTAVA, Abhishek; LOCKE, Edwin, y BARTOL, Kathryn M. (2001), "Money and subjective well-being: It's not the money, it's the motives", *Journal of Personality and Social Psychology*, vol. 80, nº 6. pp. 959-971.
897. STANLEY, James, GOWEN, Emma, y MIALL, R. Chris (2007), "Effects of agency on movement interference during observation of a moving dot stimulus", *Journal of Experimental Psychology: Human Perception and Performance*, vol. 33, nº 4. pp. 915-926.
898. STARK, Rodney (2001), *El auge del cristianismo*, traducción de Sergio Coddou, Editorial Andrés Bello, Barcelona.
899. STEIN, Edith (2004[1917]), *Sobre el problema de la empatía*, Prefacio, traducción y notas de José Luis Caballero Bono, Editorial Trotta, Madrid.
900. STEVENS, Jeffrey R., y STEPHENS, David W. (2004), "The economic basis of cooperation: Tradeoffs between selfishness and generosity", *Behavioral Ecology*, vol. 15, nº 2. pp. 255-261
901. STIGLER, George J. (1979[1965]), *Historia del pensamiento económico*, Editorial El Ateneo, Buenos Aires.
902. STONE, V. E. (2000), "The role of frontal lobes and the amygdala in theory of mind", en BARON-COHEN, S., COHEN, D.J., y TAGER-FLUSBERG, H. (eds.), *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience*, 2nd edition, Oxford University Press, Oxford. pp. 253-273.

903. STONE, V. E., COSMIDES, L., TOOBY, J., KROLL, N., Y KNIGHT, R.T. (2002), "Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage", *Proceedings of the National Academy of Sciences*, vol. 99, nº 17. pp. 11.531-11.535.
904. STOUT, Lynn A. (2008), "Taking conscience seriously", en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in economy*, Princeton University Press, Princeton. pp. 157-171.
905. STRONG, Donald R. (Jr.); y RAY, Thomas S. (Jr.) (1975), "Host tree location of a tropical vine (*Monstrora gigantea*) by skototropism", *Science*, vol. 190, nº 4.216. pp. 804-806.
906. SUDDENDORF, Thomas, y WHITEN, Andrew (2001), "Mental evolution and development: Evidence for secondary representation in children, great apes, and other animals", *Psychological Bulletin*, vol. 127, nº 5. pp. 629-650.
907. SUDGEN, Robert (1993), "Thinking as a team: towards a explanation of non-selfish behavior", *Social Philosophy and Policy*, vol. 10, nº 1. pp. 69-89.
908. SUDGEN, Robert (2000), "Team preferences", *Economics and Philosophy*, vol. 16, nº 2. pp. 175-204.
909. SUGDEN, Robert (2002), "Beyond sympathy and empathy: Adam Smit's concept of fellow-feeling", *Economics and Philosophy*, vol. 18, nº 1. pp. 63-87.
910. SUGDEN, Robert (2005), "Fellow-feeling", en GUI, Benedetto; y SUGDEN, Robert (eds.), *Economics and social interaction*, Cambridge University Press, Cambridge. pp. 52-75.
911. SUGDEN, Robert (2009), "Neither self-interest nor self-sacrifice: The fraternal morality of market relationships", en LEVIN, Simon A. (ed.), *Games, groups, and the global good*, Springer Series in Game Theory (part 3), Springer, Dordrecht (Holland). pp. 259-283.
912. SUGIYAMA, Laurece S., TOOBY, John, y COSMIDES, Leda (2002), "Cross-cultural evidence of cognitive adaptations for social exchange among the Shiwiar of Ecuadorian Amazonia", *Proceedings of the National Academy of Sciences*, vol. 99, nº 17. pp. 11.537-11.542.
913. SUROWIECKI, James (2005), *Cien mejor que uno. La sabiduría de la multitud, o por qué la mayoría siempre es más inteligente que la minoría*, traducción de J. A. Bravo, Ediciones Urano, Barcelona.
914. TAJFEL, Henry (1972), "La catégorization sociale", en MOSCOVICI, S. (ed.), *Introduction à la psychologie sociale*, vol. 1, Larousse, Paris. pp. 272-302.
915. TAJFEL, Henry (1978), "Social categorization, social identity and social comparison", en TAJFEL, Henry (ed.), *Differentiation between social*

- groups: Studies in the social psychology of intergroups relations*, Academic Press, London. pp. 27-60.
916. TAJFEL, Henry, FLAMENT, Claude, BILLIG, M.G., y BUNDY, R. F. (1971), "Social categorization and intergroup behavior", *European Journal of Social Psychology*, vol. 1, nº 2. pp. 149-178.
917. TANNER, Robin J., FERRARO, Roselina, CHARTRAND, Tanya L., BETTMAN, James R., y Van BAAREN, Rick (2007), "Of chameleons and consumption: The impact of mimicry on choice and preferences", *Journal of Consumer Research*, vol. 34, nº 6. pp. 754-766.
918. TAYLOR, Peter D.; y DAY, Troy (2004), "Cooperate with the neighbour?", *Nature*, vol. 428, nº 6.983. pp. 611-612.
919. TAYLOR, Shelley E. (2002), *Lazos vitales. De cómo el cuidado y el afecto son esenciales para nuestras vidas*, traducción de Carmen Martínez Gimeno, Taurus, Madrid.
920. TEN BRINKE, Leanne; PORTER, Stephen, y BAKER, Alysha (2012), "Darwin the detective: Observable facial muscle contractions reveal emotional high-stakes lies", *Evolution and Human Behavior*, vol. 33, nº 2. pp. 411-416.
921. TENENTI, Alberto (1999), *De las revueltas a las revoluciones*, traducción castellana de María Pons, Editorial Crítica, Barcelona.
922. THALER, Richard H. (1992), *The winner's curse: Paradoxes and anomalies of economic life*, Princeton University Press, Princeton (NJ).
923. THALER, Richard H. (2000), "From Homo Economicus to Homo Sapiens", *Journal of Economic Perspectives*, vol. 14, nº 1. pp. 133-141.
924. THOMAS, Kerry (1996), "The defensive self: A psychodynamic perspective", en STEVENS, Richard (ed.), *Understanding the self*, SAGE Publications, Thousand Oaks (CA), pp. 281-337.
925. THOREAU, Henry D. (1999[1854]), *Walden. La vida en los bosques*, traducción de Jorge Lobato, ERREPAR, Buenos Aires.
926. THORNDIKE, Edward L. (1911), *Animal intelligence. Experimental studies*, Macmillan, New York.
927. THORNHILL, Nancy W.; MARYANSKI, Alexandra M.; TOOBY, John; COSMIDES, Leda; MEYER, Peter; y TURNER, Jonathan H. (1997), "Evolutionary theory and human social institutions: Psychological foundations", en WEINGART, Peter; RICHERSON, Peter J.; MITCHELL, Sandra D.; y MAASEN, Sabine (eds.), *Human by nature. Between biology and the social sciences*, Lawrence Erlbaum Associates Publishers, Mahwah (New Jersey) & London. pp. 201-252.
928. TIERNEY, Briam (1997), *The idea of natural rights: Studies on natural rights, natural law, and church law. 1150-1625*, B. Berdmans, Michigan.

929. TIROLE, Jean (1996), "A Theory of collective reputations (with applications to the persistence of corruption and to firm quality)", *Review of Economic Studies*, vol. 63, nº 1. pp. 1-22.
930. TOBEÑA, Adolf (2005), *Mártires mortíferos: Un itinerario por el cerebro de los suicidas atacantes*, Universitat de Valencia. Servei de Publicacions, Valencia.
931. TODOROV, Alexander, BOBBINI, Ida M., EVANS, Karla, y HAXBY, James V. (2007), "Spontaneous retrieval of affective person knowledge in face perception", *Nuropsychologia*, vol. 45, nº 1. pp. 163-173.
932. TOM, Sabrina (2007), "The neural basis of loss aversion in decision-making under risk", *Science*, vol. 315, nº 5.811. pp. 515-518.
933. TOMASELLO, Michael, y VAISH, Amrisha (2013), "Origins of cooperation and morality", *Annual Review of Psychology*, vol. 64, nº 1. pp. 231-255.
934. TOMASELLO, Michael; CARPENTER, Malinda; CALL, Josep; BEHNE, Tanya; y MOLL, Henrike (2005), "Understanding and sharing intentions: The origins of cultural cognition", *Behavioral and Brain Sciences*, vol. 28, nº 5. pp. 675-735.
935. TOMASELLO, Michael; MELIS, Alicia P.; TENNIE, Claudio; WYMAN, Emily, y HERRMANN, Esther (2012), "Two key steps in the evolution of human cooperation. The interdependence hypothesis", *Current Anthropology*, vol. 53, nº 6. pp. 673-692.
936. TÖNNIES, Ferdinand (1979[1887]), *Comunidad y asociación*, traducción de José-Francisco Ivars, prólogo de Salvador Giner y Lluís Flaquer, y colaboración técnica de Salvador Giner, Ediciones Península, Barcelona.
937. TONONI, Giulio; y EDELMAN, Gerard M. (1998), "Consciousness and Complexity", *Science's Compass Review*, vol. 282, nº 5.395. pp. 1.846-1.851.
938. TOOBY, John; COSMIDES, Leda; y PRICE, Michael E. (2006), "Cognitive adaptations for *n*-persons exchange: The evolutionary roots of organizational behavior", *Managerial and Decision Economics*, vol. 27, nº 2-3. pp. 103-129.
939. TOOBY, John; y COSMIDES, Leda (1996), "Friendship and the banker's paradox: Other pathways to the evolution of adaptations for altruism", *Proceedings of the British Academy*, vol. 88, nº 1. pp. 119-143.
940. TOOBY, John; y COSMIDES, Leda (2007), "Evolutionary psychology, ecological rationality, and the unification of the behavioral sciences", *Behavioral and Brain Sciences*, vol. 30, nº 1. pp. 42-43.

941. TORO IBÁÑEZ, Miguel Ángel, y CASTRO NOGUEIRA, Laureano (1999), “Modelos sociobiológicos de la evolución del altruismo”, *Anthropos*, nº 182. pp. 42-48.
942. TRAULSEN, Arne, SEMMANN, Dirk, SOMMERFIELD, Ralf D., KRAMBECK, Hans-Jürgen, y MILINSKY, Manfred (2010), “Human strategy updating in evolutionary games”, *Proceedings of the National Academy of Sciences*, vol. 107, nº 7. pp. 2.962-2.966.
943. TRAULSEN, Arne, y NOWAK, Martin A. (2006), “Evolution of cooperation by multilevel selection”, *Proceedings of the National Academy of Sciences*, vol. 103, nº 29. pp. 10.952-10.955.
944. TREMBLAY, León; y SCHULTZ, Wolfram (1999), “Relative reward preference in primate orbitofrontal cortex”, *Nature*, vol. 398, nº 6.729. pp. 704-708.
945. TRICOMI, E., RANGEL, A., CAMERER, C., y O'DOHERTY, J. (2010), “Neural evidence of inequality-averse social preferences”, *Nature*, vol. 463, nº 7.284. pp. 1.089-1.091.
946. TRINCADO, Estrella (2004), “Equity, utility and transaction costs: On the origin of judicial power in Adam Smith”, *Storia del Pensiero Economico*, Nuova Serie, vol. 1, nº 1. pp. 33-51.
947. TRIVERS, Robert L. (1972), “Parental investment and sexual selection”, en CAMPBELL, Bernard G. (ed.), *Sexual selection and the descent of man*, Aldine de Gruyter, Chicago. pp. 136-139.
948. TRIVERS, Robert L. (1978[1971]), “The evolution of reciprocal altruism”, (*Quarterly Review of Biology*, vol. 46, nº 1. pp. 35-57). Reimpreso en CLUTTON-BROCK, T. H., y HARVEY, Paul H. (eds.) (1978), *Readings in Sociobiology*, W. H. Freeman & Company, San Francisco (California). pp. 189-226.
949. TRIVERS, Robert L. (2000), “The elements of a scientific theory of self-deception”, *Annals of the New York Academy of Sciences*, vol. 907. pp. 114-131.
950. TRIVERS, Robert L. (2004), “Mutual benefits at all levels of life”, *Science*, vol. 304, nº 5.673. pp. 964-965.
951. TULLOCK, Gordon (1971), “The paradox of revolution”, *Public Choice*, vol. 9, otoño. pp. 89-99.
952. TULLOCK, Gordon (1987), “Biological applications of economics”, en EATWELL, John; MILGATE, Murray; y NEWMAN, Peter (eds.), *The New Palgrave. A Dictionary of Economics*, vol. I, Macmillan, London. pp. 246-247.
953. TULLOCK, Gordon (1990), “The economics of (very) primitive societies”, *Journal of Social and Biological Structures*, vol. 13, nº 2. pp. 151-162.

954. TUOMELA, Raimo (2000), *Cooperation: A philosophical study*, Kluwer Academic Publishers, Dordrecht & Boston.
955. TUOMELA, Raimo (2007), *The philosophy of sociality: The shared point of view*, Oxford University Press, Oxford.
956. TURCHET, Philippe (2010), *El lenguaje del cuerpo*, traducción de José Antonio Aguirre Setas, Ediciones Mensajero, Bilbao.
957. TURNER, Jonathan H. (2000), *On the origins of human emotions. A sociological inquiry into the evolution of human affect*, Stanford University Press, Stanford.
958. TVERSKY, Amos; y KAHNEMAN, Daniel (1982), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.
959. TVERSKY, Amos; y KAHNEMAN, Daniel (1987), "Rational choice and the framing of decisions"; en HOGARTH, R. M., y REDER, M. W. (eds.), *Rational choice: The contrast between economics and psychology*, University of Chicago Press, Chicago. pp. 67-94.
960. TYLER, Tom R., y BLADER, Steven L. (2000), *Cooperation in groups. Procedural justice, social identity, and behavioral engagement*, Psychology Press, New York.
961. TYLER, Tom R., y BLADER, Steven L. (2001), "Identity and cooperative behavior in groups", *Group Processes & Intergroup Relations*, vol. 4, nº 3. pp. 207-226.
962. TYLER, Tom R., y BLADER, Steven L. (2003), "The group engagement model: Procedural justice, social identity, and cooperative behavior", *Personality and Social Psychology Review*, vol. 7, nº 2. pp. 349-361.
963. URGESI, Cosimo, CANDIDI, Matteo, IONTA, Silvio, y AGLIOTI, Salvatore M. (2007), "Representation of body identity and body actions in extrastriate body area and ventral premotor cortex", *Nature Neuroscience*, vol. 10, nº 1. pp. 30-31.
964. URRUTIA ELEJALDE, Juan (2014), "A la individuación por la pertenencia", *Energeia*, vol. 6, nº 1. pp. 43-60.
965. URRY, Heather L., NITSCHKE, Jack B., DOLSKI, Isa, JACKSON, Daren C., DALTON, Kim M., MUELLER, Corrina J., ROSENKRANZ, Melissa A., RYFF, Carol D., SINGER, Burton H., y DAVIDSON, Richard J. (2004), "Making a life worth living. Neural correlates of well-being", *Psychological Science*, vol. 15, nº 6. pp. 367-372.
966. Van LANGE, Paul A. M.; LIEBRAND, Wim B.; MESSICK, David M.; y WILKE, Henk A. (1992), "Social dilemmas: The state of the art", en LIEBRAND, W. B.; MESSICK, D. M., y WILKE, H. A. (eds.), *Social dilemmas: Theoretical issues and research findings*, Pergamon Press, London. pp. 3-28.

967. VEBLEN, Thorstein (2002[1899]), *Teoría de la clase ociosa*, con introducción de John Kenneth Galbraith, traducción de Vicente Herreros, Fondo de Cultura Económica, México.
968. VERBALIS, J. G. (1999), “The brain oxytocin receptors”, *Frontiers in Neuroendocrinology*, vol. 20, n° 2. pp. 146-156.
969. VIANELLO, Michelangelo, GALLIANI, Elisa Maria, y HAIDT, Jonathan (2010), “Elevation at work: The effects of leaders' moral excellence”, *Journal of Positive Psychology*, vol. 5, n° 5. pp. 390-411.
970. VIGNEMONT, Frederique (2006), “When do we empathize?”, en BOCK, Greg, y GOODE, Jamie (eds.), *Empaty and fairness*, Novartis Foundation, London. pp. 181-196.
971. VIGNEMONT, Frederique; y SINGER, Tania (2006), “The empathic brain: How, when and why?”, *Trends in Cognitive Sciences*, vol. 10, n° 10. pp. 435-441.
972. VUILLEUMIER, P., ARMONY, J.L., CLARKE, K., DRIVER, J., DOLAN, R.J. (2002), “Neural response to emotional faces with and without awareness”, *Neuropsychologia*, vo. 40. pp. 2.156-2.166.
973. WADE, Nicholas (2015), *Una herencia incómoda. Genes, raza e historia humana*, traducción de Juandomènec Ros, Editorial Ariel, Barcelona.
974. WARNEKEN, Felix, y TOMASELLO, Michael (2006), “Altruistic helping in human infants and young chimpanzees”, *Science*, vol. 311, n° 5.765. pp. 1.301-1.303.
975. WATSON, John (1966), *Teorías del placer*, Editorial Paidós, Buenos Aires.
976. WATSON, Peter (2012), *The great divide: History and human nature in the Old world and the New*, Harper Collins Publishers, New York.
977. WEBER, J. Mark, y MURNIGHAN, J. Keith (2008), “Suckers or saviors? Consistent contributors in social dilemmas”, *Journal of Personality and Social Psychology*, vol. 95, n° 6 p. 1.340-1.353.
978. WEBER, Max (1969[1922]), *Economía y sociedad. Esbozo de sociología comprensiva*, Vol. I, primera reimpresión en español de la segunda edición española de la cuarta edición en alemán, Fondo de Cultura Económica, México.
979. WEDEKIND, Claus, SEEBECK, Thomas, PAEPKE, Alexander J. (1995), “MHC-dependent mate preferences in humans”, *Proceedings of the Royal Society of London, Biological Sciences*, vol. 260, n° 1.359. pp. 245-249.
980. WEGNER, Daniel M. (2003), “The mind's self-portrait”, *Annals of the New York Academy of Sciences*, vol. 1.001. pp. 212-225.

981. WEIMANN, Joachim (1994), "Individual behavior in a free riding experiment", *Journal of Public Economics*, vol. 54, nº 1. pp. 185-200.
982. WEISMANN, Eberhard (1994), *Los rituales amorosos. Un aspecto fundamental en la comunicación de los animales*, traducción de la edición alemana de Otro Maier Verlag por Paula Álvarez, Salvat Editores, Barcelona.
983. WENG, Helen Y., FOX, Andrew S., SHACKMAN, Alexander J., STODOLA, Diane, E., CALDWELL, Jessica Z. K., OLSON, Matthew C., ROGERS, Gregory M., y DAVIDSON, Richard J. (2013), "Compassion training alters altruism and neural responses to suffering", *Psychological Science*, vol. 24, nº 7. pp. 1.171-1.180.
984. WESNER, Erik (2010), *Success made simple: An inside look at why amish businesses thrive*, Wiley, San Francisco (CA).
985. WEST, Stuart A.; PEN, Ido; y GRIFFIN, Ashleigh S. (2002), "Cooperation and competition between relatives", *Science*, vol. 296, nº 5.565. pp. 72-75.
986. WHEELER, Michael, y Clark, Andy (2008), "Culture, embodiment and genes: unraveling the triple helix", *Philosophical Transactions of the Royal Society, Biological Sciences*, vol. 363, nº 1.509. pp. 3.563-3.575.
987. WHITMEYER, Joseph M. (1997), "Endogamy as a basis for ethnic behavior", *Sociological Theory*, vol. 15, nº 1. pp. 162-178.
988. WICHMAN, Harvey (1970), "Effects of isolation and communication on cooperation in a two-person game", *Journal of Personality and Social Psychology*, vol. 16, nº 1. pp. 114-120.
989. WICKSTEED, Philip H. (1963[1914]), "Alcance y método de la economía política", en STIGLER, George J., y BOULDING, Kenneth E., (dir.), *Ensayos sobre teoría de los precios*, Ediciones Aguilar, Madrid. Reimpresión de *The Economic Journal*, vol. XXIV, 1914. pp. 1-23.
990. WILBER, Ken (2000), *Una teoría de todo. Una visión integral de la ciencia, la política, la empresa y la espiritualidad*, Editorial Kairós, Barcelona.
991. WILKINSON, Richard (2001), *Las desigualdades perjudican. Jerarquías, salud y evolución humana*, traducción de Silvia Furió, Editorial Crítica, Barcelona.
992. WILLIAMS, George C. (1996[1966]), *Adaptation and natural selection: A critique of some current evolutionary thought*, reedición con nuevo prefacio del autor, Princeton University Press, Princeton.
993. WILLIAMSON, Oliver E. (1989), *Las instituciones económicas del capitalismo*, Fondo de Cultura Económica, México.

994. WILLIAMSON, Oliver E. (1993), "Calculativeness, trust, and economic organization", *Journal of Law and Economics*, vol. 36, n° 1. pp. 453-502.
995. WILSON, David S.; y SOBER, Elliott (1994), "Reintroducing group selection to the human behavioral sciences", *Behavioral and Brain Sciences*, vol. 17, n° 4. pp. 585-654.
996. WILSON, David Sloan (1983), "The group selection controversy: History and current status", *Annual Review of Ecology Systematics*, vol. 14. pp. 159-187.
997. WILSON, David Sloan (2002), *Darwin's cathedral: Evolution, religion, and nature of society*, University of Chicago Press, Chicago.
998. WILSON, David Sloan (2007), *Evolution for everyone: How Darwin's theory can change the way we think about our lives*, Delta, New York.
999. WILSON, Edward O. (1980[1975]), *Sociobiología: La nueva síntesis*, traducción de Ramón Navarro, revisado por Andrés de Haro, Editorial Omega, Barcelona.
1000. WILSON, Edward O. (1983), *Sobre la naturaleza humana*, traducción al castellano de A. M. Sánchez, Fondo de Cultura Económica, México.
1001. WILSON, Edward O. (1998), *Consilience. The unity of knowledge*, Little Brown & Co., London.
1002. WILSON, Edward O. (2005), "No somos un superorganismo", entrevista publicada en PUNSET, Eduardo, *Cara a cara con la vida, la mente y el Universo. Conversaciones con los grandes científicos de nuestro tiempo*, Círculo de Lectores, Barcelona.
1003. WILSON, Edward O. (2012), *La conquista social de la Tierra*, traducción de Juandomèmec Ros I Aragonés, Editorial Debate, Barcelona.
1004. WILSON, Edward O.; y HÖLLDOBLER, Bert (2005), "Eusociality: Origin and consequences", *Proceedings of the National Academy of Sciences USA*, vol. 102, n° 38. pp. 13.367-13.371.
1005. WILSON, James (1993), *The moral sense*, The Free Press, New York.
1006. WILSON, Margo, y DALY, Martin (1997), "Adaptations for social living: Relationship-specific social psychological adaptations", en BOCK, G., y CARDEW, G. (eds.), *CIBA Foundation Symposium on Characterizing Psychological Adaptations*, Wiley, Chichester (UK).
1007. WILSON, Margo, y DALY, Martin (1997), "Relationship-specific social psychological adaptations", en BOCK, G., Y CARDEW, G. (eds.), *CIBA Foundation Symposium on Characterizing Psychological Adaptations*, Wiley, Chichester. pp. 253-268.

1008. WILSON, Robert A.; y KEIL, Frank (eds.) (2002), *Enciclopedia MIT de Ciencias Cognitivas*, vol. I, edición en español de la primera reimposición en inglés, Editorial Síntesis, Madrid.
1009. WILTERMUTH, Scott S., y HEATH, Chip (2009), “Synchrony and cooperation”, *Psychological Science*, vol. 20, nº 1. pp. 1-5.
1010. WISPÉ, Lauren (1976), “Simpatía y empatía”, en SILLS, David L. (dir.), *Enciclopedia Internacional de las Ciencias Sociales*, vol. 9, Editorial Aguilar, Madrid. pp. 619-623.
1011. WISPÉ, Lauren (1991), *The psychology of sympathy*, Plenum Press, New York & London.
1012. WITTGENSTEIN, Ludwig J. (1999[1921]), *Tractatus logico-philosophicus*, traducción e introducción de Jacobo Muñoz e Isidoro Reguera, Alianza Editorial, Madrid, 1997.
1013. WOHLSCHLAGER, A.; HAGGARD, P.; GESIERICH, B., y PRINZ, W. (2003), “The perceived onset time of self -and other- generated actions”, *Psychology Science*, vol. 14, nº 6. pp. 586-591.
1014. WYNNE-EDWARDS, Vero C. (1962), *Animal dispersion in relation to social behavior*, Hafner Publishing, New York.
1015. WYNNE-EDWARDS, Vero C. (1986), *Evolution through group selection*, Blackwell, Oxford.
1016. XU, Xiaojing; ZUO, Xiangyu; WANG, Xiaoying; y HAN, Shihui (2009), “Do you feel my pain? Racial group membership modulates empathic neural responses”, *Journal of Neuroscience*, vol. 29, nº 26. pp. 8.525-8.529.
1017. YAMAGISHI, Toshio (1986), “The provision of a sanctioning system as a public good”, *Journal of Personality and Social Psychology*, vol. 51, nº 1. pp. 110-116.
1018. YAMAGISHI, Toshio; KANAZAWA, Satoshi; MASHIMA, Rie; y TERAJ, Shigeru (2005), “Separating trust from cooperation in a dynamic relationship. Prisoner's dilemma with variable dependence”, *Rationality and Society*, vol. 17, nº 3. pp. 275-308.
1019. YEUNG, N., y SANFEY, A. G. (2004), “Independent coding of reward magnitude and valence in the human brain”, *Journal of Neuroscience*, vol. 24, nº 28. pp. 6258-6264.
1020. YOUNG, Liane, CUSHMAN; Fiery, HAUSER, Mark, SUXE, Rebecca (2007), “The neural basis of the interaction between theory of mind and moral judgement”, *Proceedings of the National Academy os Sciences*, vol. 104, nº 20. pp. 8.235-8.240.

1021. ZAJONC, Robert B. (1968), "Attitudinal effects of mere exposure", *Journal of Personality and Social Psychology*, Monograph Supplement, vol. 9, n° 2 (part 2). pp. 1-27.
1022. ZAJONC, Robert B. (1980), "Feeling and thinking: preferences need no inference", *American Psychologist*, vol. 35, n° 2. pp. 151-175.
1023. ZAK, Paul J. (2005), "Trust: A temporary human attachment facilitated by oxytocin", *Behavioral and Brain Sciences*, vol. 28, n° 3. pp. 368-369.
1024. ZAK, Paul J. (2008a), "Values and value: Moral economics", en ZAK, Paul J. (ed.), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton. pp. 259-278.
1025. ZAK, Paul J. (ed.) (2008b), *Moral markets: The critical role of values in the economy*, Princeton University Press, Princeton.
1026. ZAK, Paul J.; STANTON, Angela A.; y AHMADI, Sheila (2007), "Oxytocin increases generosity in humans", *PlosONE*, vol. 2, n° 11. pp. 1-5. Edición en hipertexto: <http://dx.doi.org/10.1371/journal.pone.0001128>
1027. ZAMAGNI, Stefano (1996), *Economics of altruism*, Edward Elgar Publishing, Aldershot (UK).
1028. ZEAMAN, David J. (1949), "Response latency as a function of the amount of reinforcement", *Journal of Experimental Psychology*, vol. 39, n° 4. pp. 466-483.
1029. ZEGGELINK, Evelien P. H., De VOS, Henk, y ELSAS, Donald (2000), "Reciprocal altruism and group formation: The degree of segmentation or reciprocal altruists who prefer 'old helping partners'", *Journal of Artificial Societies and Social Simulation*, vol. 3, n° 3. Edición en hipertexto: <http://jasss.soc.surrey.ac.uk/3/3/1.html>
1030. ZUCKER, Lynn G. (1986), "Production of trust: Institutional sources of economic structure. 1840-1920", *Research in Organizational Behavior*, vol. 8, n° 1. pp. 53-111.
1031. ZWIRN, Gregor (2007), "Methodological individualism or methodological atomism: The case of Friedrich Hayek", *History of Political Economy*, vol. 39, n° 1. pp. 47-80.