

<http://doi.org/10.21555/top.v62i0.1233>

The Evolution of Moral Judgment. A Plural, Functional Explanation

La evolución del juicio moral. Una explicación funcional plural

Víctor Emilio Parra Leal
Universidad Nacional de Colombia
Colombia
vparral@unal.edu.co
<https://orcid.org/0000-0002-1874-4927>

Recibido: 09 – 09 – 2019.
Aceptado: 10 – 01 – 2020.
Publicado en línea: 10 – 12 – 2021.



This work is licensed under a Creative Commons Attribution
-NonCommercial-ShareAlike 4.0 International License.

Abstract

Evolutionary functional theories about morality assume that morality has been evolutionarily selected because it produces a cooperative and pro-social effect (*selected effect* functional theory, SE). This assumption wrongly implies that the trait which produces the selected effect is fully known. Furthermore, a SE theory does not seem to explain the peculiarities of moral thought, which are mostly synthesized in the features of moral judgment: its authoritative and inescapable character (practical clout). This paper provides an alternative SE explanation and argues that a *causal role* functional explanation (CR) is also necessary because this kind of explanation establishes the way the selected effect interacts with basic human evaluative propensities, which not only makes it possible to give an account of the practical clout of moral judgment, but of the phenomenon of moral ambivalence as well. However, the connection between reason and emotion and their influence on morality remains unexplained.

Keywords: morality; evolution of morality; functional theories; practical clout; moral judgment.

Resumen

Las teorías funcionales evolutivas de la moral asumen que la moral ha sido seleccionada evolutivamente por su efecto de lograr un comportamiento cooperativo y prosocial (explicación funcional del *efecto seleccionado*, ES). Esta asunción implica incorrectamente que se conoce el rasgo o rasgos que conforman la moral. Adicionalmente, la teoría funcional ES no parece dar cuenta de las particularidades del pensamiento moral, las cuales se sintetizan principalmente en las características que se le atribuyen al juicio moral, a saber, su autoridad e inevitabilidad (fuerza práctica). Este artículo plantea una explicación ES alternativa y argumenta que también se requiere de una explicación funcional del *rol causal* (RC), la cual muestra cómo se integra el efecto seleccionado con ciertas tendencias evaluativas básicas humanas, lo que permite dar cuenta de la fuerza práctica del juicio moral y de fenómenos como la ambivalencia moral. Queda sin explicar cuál es el rol que juegan la razón y las emociones en la formulación del juicio moral.

Palabras clave: moral; evolución de la moral; teorías funcionales; fuerza práctica; juicio moral.

1. Introducción

El debate funcional en biología ha estado caracterizado por la oposición entre dos posturas predominantes: la postura etiológica o del efecto seleccionado (ES) (Wright, 1973; Millikan, 1984; Neander, 1991), la cual considera que la función de un rasgo está determinada por su origen, es decir, por la razón por la cual existe dicho rasgo, y la postura organizacional o del rol causal (RC) (Cummins, 1975; Davies, 2009; Cummins y Roth, 2010), que sostiene que no es el origen del rasgo, sino el papel que dicho rasgo juega dentro de la consecución de un objetivo que beneficia o aporta a un sistema al cual pertenece el rasgo lo que determina la atribución funcional.

La postura etiológica es capaz de dar cuenta de la diferencia entre la función de un determinado rasgo y sus efectos puramente accidentales, de suerte que es capaz de distinguir, por ejemplo, entre el efecto esencial que produce el corazón de bombear sangre y el “efecto accidental” de producir un sonido relativamente rítmico; la teoría organizacional, por su parte, permite adscribir funciones sin necesidad de conocer el origen o la historia de determinado rasgo. La función de un hueso del cuerpo, por ejemplo, está determinada por el rol que cumple dicho elemento en la articulación con otros huesos para darle estabilidad y protección al cuerpo, lo cual no se determina necesariamente conociendo el origen del rasgo, sino observando el rol organizacional que desempeña dicho rasgo dentro del sistema.

Esta aparente oposición ha alcanzado una síntesis en teorías funcionales plurales que, a partir de evaluar cómo se usa el concepto de “función biológica” en la práctica científica real, sostienen que son necesarias explicaciones tanto ES como RC (Amundson y Lauder, 1994; Huneman, 2013; Jackson y Fry, 2016).

Frente a esta variedad de teorías funcionales, es válido preguntarse cuál de todas puede dar cuenta de mejor manera de la forma en que se usa el concepto de “función biológica” en las teorías funcionales evolutivas de la moral. Las teorías funcionales biológicas de la moral comparten de manera implícita ciertas asunciones, lo cual tiene implicaciones interesantes respecto a la validez de dichas teorías y respecto a qué requiere una teoría funcional biológica de la moral para ser convincente. La concepción funcional adoptada en las teorías funcionales de la moral no ha sido explicitada y hacerlo es uno de los objetivos de este artículo.

Una vez hecho esto, será necesario establecer a qué retos se enfrenta tal explicación e intentar ofrecer una alternativa que pueda abordarlos.

2. La concepción funcional implícita en las teorías funcionales evolutivas de la moral y los retos a los que se enfrenta

Muchos filósofos coinciden en que la moral humana es el rasgo, o conjunto de rasgos, que tienen la función de reducir las tensiones sociales y asegurar una convivencia pacífica entre los miembros de una sociedad.

Por ejemplo, Prinz (2007) sostiene que la moralidad emerge como un conjunto de reglas con la función de hacer que la gente trabaje colectivamente y de una manera estable y productiva, mientras que Joyce (2006) cree que el pensamiento moral tiene la función de contribuir a la satisfacción de los intereses a largo plazo y hacer más efectiva la negociación colectiva al permitir el castigo y al unir a los individuos en una red compartida de toma de decisiones.

Otros, como Sinclair (2012), atribuyen al juicio moral la función de lograr una coordinación interpersonal y conciben a los juicios morales como producto de un mecanismo que permite la existencia de grupos de individuos que coordinan sus acciones y emociones para un beneficio mutuo. Bajo esta perspectiva, la función de la moral es producir patrones mutuamente beneficiosos de acción y emoción.

Finalmente, Kitcher (2006) considera que la función de la moral es habilitar la cohesión social a través de la amplificación de las disposiciones altruistas del ser humano. Estas definiciones funcionales, en lo fundamental, asumen que la moral humana es producto de la evolución biológica, específicamente de la evolución por selección natural (Darwin, [1882] 1989, cap. IV).¹ Es decir, los rasgos fundamentales de la moral humana se conciben como producto de la evolución biológica.

Pero las definiciones dadas ejemplifican una dificultad presente en las teorías funcionales biológicas de la moral en general, a saber, la falta de acuerdo respecto a cuál es el rasgo o conjunto de rasgos que han sido seleccionados para cumplir la función de hacer al ser humano prosocial y cooperativo. Mientras que Prinz habla de “un conjunto de reglas”, Kitcher menciona un conjunto de “disposiciones”, y Sinclair se

¹ La selección natural (SN) es el mecanismo evolutivo que consiste en la reproducción diferencial de los genotipos de una población biológica.

refiere a la moral de manera más delimitada, haciendo referencia a un "mecanismo".

Esta diferencia marca una tendencia en las teorías funcionales de la moral entre quienes, siguiendo a Darwin ([1882] 1989, p. 106), hacen referencia a un "órgano, mecanismo o sentido moral" (Joyce, 2006, p. 57; cfr. Peterson, 2011, p. 58 y Sinclair, 2012, p. 14) y quienes hacen referencia a la moral como un conjunto de propiedades, bien sean disposiciones, emociones, actitudes, etc. (cfr. Flack y De Waal, 2002, p. 23; Bekoff y Pierce, 2009, p. 7).

La diferencia es aparentemente superficial, pues puede argumentarse que quienes hablan de un sentido, órgano o mecanismo moral están simplemente etiquetando al complejo conjunto de actitudes, emociones y mecanismos que intervienen en la producción del efecto cooperativo bajo un único concepto o término. El mismo Darwin hace referencia a la moral muchas veces como un conjunto de disposiciones y actitudes complejas que involucran tanto aspectos biológicos como culturales ([1882] 1989, p. 114). Sin embargo, la tendencia a unificar los diversos elementos de la moral bajo un mecanismo claramente delimitable responde a la naturaleza misma de la explicación funcional adoptada.

Como ya se dijo, se asume que la moral (sea cual sea el rasgo o rasgos que la conformen) ha sido seleccionada por generar comportamientos prosociales y cooperativos (Darwin, [1882] 1989, p. 119), lo cual significa que las explicaciones funcionales evolutivas de la moral implícitamente asumen que la teoría funcional biológica que mejor se adecúa a ellas es la etiológica o del *efecto seleccionado* (ES).

Según la teoría funcional ES, la función de un rasgo biológico consiste en el efecto que produce dicho rasgo, por el cual ha sido seleccionado. Eso significa que los comportamientos cooperativos y prosociales constituyen la función de la moral. Sin embargo, sin identificar el (o los) rasgo(s) que cumple(n) tal función, la atribución funcional se diluye entre diversos elementos, sin claridad respecto a la relación adecuada que han de tener entre sí para que la moral cumpla adecuadamente su función, o sin un consenso respecto al papel y jerarquía de dichos elementos en el cumplimiento de la función de la moral. Esto hace que las teorías funcionales disponibles pongan en elementos distintos la importancia del cumplimiento de la función de la moral.

Suelen postularse dos capacidades humanas que evolucionaron y que, unidas, pueden explicar la moral o, más específicamente, la cognición moral. Tales capacidades consisten en un sistema basado en

impulsos primitivos humanos que explican los sentimientos empáticos del ser humanos hacia sus semejantes, y un sistema cognitivo, representado por la capacidad humana de plasmar en el lenguaje las normas e imperativos morales (cfr. Gibbard, 1990, pp. 56-57).

Darwin también sostenía esta concepción dual, pues concebía al “sentido moral” como una mezcla de “instintos sociales” y “poderes intelectuales”. Dentro de los poderes intelectuales a los que se refiere Darwin están la memoria, el lenguaje y el hábito, capacidades que no son exclusivas del dominio moral (cfr. Darwin, [1882] 1989, p. 70).

El hecho de que en la moral se incluyan capacidades multipropósito, que no son exclusivas del dominio moral, pone en riesgo la idea de que existe un rasgo exclusivo que produce el efecto de hacernos prosociales y cooperativos² y que es justamente por producir este efecto que ha sido seleccionado.

La primera de tales posturas es la de quienes creen que la moral es un subproducto de la conjunción de adaptaciones que evolucionaron con propósitos inicialmente distintos al de la moral pero que, al unirse, dan cuenta apropiadamente de la moral (cfr. Ayala, 2010, p. 9015; Prinz, 2008, p. 403). Esta postura no necesariamente atenta contra la idea de que la moral es un rasgo definido seleccionado, pues es muy frecuente en el reino biológico que elementos que inicialmente evolucionaron con propósitos distintos, una vez combinados, han sido cooptados³ para

² En este punto es necesario hacer la siguiente aclaración: se ha dicho que la moral es el rasgo que produce el efecto seleccionado de hacer al ser humano prosocial y cooperativo. Podría argumentarse que el rasgo no es lo esencial en la moral, sino que lo esencial es el efecto seleccionado que produce dicho rasgo, a saber, el comportamiento prosocial y cooperativo. Sin embargo, el rasgo es determinante en la moral humana en la medida en que la moral no es simplemente el comportamiento prosocial y cooperativo, sino una forma particular en la que se logra dicho comportamiento. Las abejas y las hormigas, por ejemplo, exhiben también comportamientos cooperativos, pero difícilmente se dice que son animales morales. Luego, el rasgo que produce de manera particular el efecto resulta determinante para investigar qué es lo propio y esencial del fenómeno moral. Como se verá, la manera particular de producir dicho efecto en el ser humano es a través del juicio moral con su correspondiente “fuerza práctica”.

³ La cooptación o exaptación es un concepto biológico que explica el frecuente fenómeno evolutivo, en el que un rasgo es seleccionado para un propósito determinado y, posteriormente (luego de millones de años), empieza

cumplir una nueva función. Este podría ser el caso de las explicaciones de la moral que apelan a una conjunción de elementos o rasgos no exclusivos de la moral.

Sin embargo, hay quienes niegan la existencia de un “sentido moral” que ha evolucionado para fortalecer la cohesión social y la cooperación, y consideran que la moral no es producto de la evolución biológica, sino que la atribuyen a procesos puramente culturales (Machery y Mallon, 2010).⁴

Por lo tanto, el principal reto de una explicación evolutiva funcional de la moral tiene que ver con mostrar cómo pudo haber evolucionado el sentido moral, bien sea entendido como la cooptación de diversos rasgos, o como un único rasgo claramente definido con la función que se le atribuye a la moral, pero fundamentalmente biológico y no producto exclusivo de la cultura humana. Pero este reto no puede abordarse si no existe primero claridad sobre lo que es la cognición moral.

Joyce defiende la existencia de un mecanismo moral exclusivo. Dicho mecanismo moral tiene como principal tarea producir juicios morales (cfr. Joyce, 2006, p. 134). Joyce se centra en el juicio moral porque cree que, si algo es innato en el ser humano respecto a la moral, es la tendencia a hacer juicios morales o a pensar en términos morales (p. 70). En ese sentido, la caracterización de la cognición moral está dada por las características del juicio moral.

Siguiendo a Joyce, han de caracterizarse tales juicios como afirmaciones que: 1) no pretenden ser consideraciones prudenciales, sino que más bien son inevitables y no dependen de los intereses o fines de aquéllos a quienes van dirigidos; 2) trascienden las convenciones humanas; 3) gobiernan las relaciones interpersonales, de suerte que

a adquirir una función diferente, la cual, al ser también adaptativa, facilita la selección del rasgo, pero con un nuevo propósito. El ejemplo clásico de exaptación es el de las plumas de las aves, las cuales inicialmente tenían una función termoreguladora y posteriormente fueron cooptadas para el vuelo (Gould y Vrba, 1982).

⁴ No es la intención de este artículo negar que la cultura posee influencia sobre la moral; sin embargo, se sostiene la tesis —que no riñe con asumir una gran influencia de la cultura en la moral— de que la evolución biológica es el principal determinante de la moral. De hecho, la explicación funcional ES tiene que, de una forma u otra, aceptar esta tesis.

están diseñadas para combatir el individualismo, y 4) implican nociones de merecimiento y justicia (cfr. 2006, p. 70).

Las características 1 y 2 que Joyce atribuye al juicio moral están relacionadas con el concepto de “deber moral” o con lo que Joyce llama la “fuerza práctica” del juicio moral. Según Joyce, el juicio moral posee fuerza práctica (*practical clout*), pues es inevitable y autoritativo. La inevitabilidad es definida por Joyce como esa obligación que se siente de cumplir los preceptos morales con independencia de los intereses, necesidades y deseos particulares que se puedan tener (característica 1). El juicio “robar es malo”, por ejemplo, parece exigir un tipo de acción que no depende de los deseos contingentes o inclinaciones particulares de los seres humanos. A pesar de que alguien se divierta robando, el juicio parece *obligarlo* a no hacerlo. Igualmente, el juicio moral prescribe de manera indirecta un determinado curso de acción que no se puede ignorar. Esta característica de los juicios morales corresponde a lo que Joyce denomina el “carácter autoritativo” del juicio moral. Un juicio moral posee autoridad si los individuos tienen ciertas razones deliberativas de peso para seguirlos. O, dicho de otra forma, el juicio posee carácter autoritativo si se puede considerar que alguien actúa irracionalmente al ignorarlo, lo cual significa que el juicio moral trasciende las convenciones humanas (característica 2) (cfr. Joyce, 2006, p. 62).

Resulta bastante simplista reducir un fenómeno tan complejo como la moral humana a los juicios morales. Sin embargo, hay dos razones por las cuales el juicio moral, tal como lo concibe Joyce, puede considerarse como la clave de la cognición moral: primero, porque da cuenta de esa clase de autoridad propia y particular de los juicios morales que se denomina “autoridad” o “deber moral” y que tanto ha intrigado a algunos filósofos (como Immanuel Kant, J. Mackie, Michael Smith y Christine Korsgaard, entre otros); en segundo lugar, porque en la producción del juicio moral —tal como la concibe Joyce— parecen sintetizarse los elementos emocionales y racionales del ser humano a los que ya se hizo referencia (Joyce, 2006).

Este artículo no busca solucionar de manera definitiva las discusiones sobre lo que es un juicio moral ni si el “deber moral” es parte esencial de este. Podría argumentarse que no todos los juicios morales poseen la fuerza práctica que les atribuye Joyce. Sin embargo, es un hecho fenomenológico que el ser humano *percibe* los juicios morales como poseyendo una clase particular de autoridad, que es ineludible (Ruse y Wilson, 1996; Joyce, 2006; Kitcher, 2006). Prueba de ello son los estudios

que muestran diferencias esenciales entre las normas morales y las de etiqueta, siendo el caso que las primeras tienden a ser percibidas como independientes de la autoridad, a estar relacionadas con el daño a los demás y a poseer, en general, una suerte de obligatoriedad de la que carecen las primeras (Smetana y Braeges, 1990).

Lo importante es tener presente que una explicación funcional ES de la moral tendrá que dar cuenta del rasgo o rasgos que producen el efecto seleccionado, así como de la cognición moral entendida como la capacidad de producir juicios con las características que les adjudica Joyce a los juicios morales, especialmente la de poseer fuerza práctica. A continuación se ofrece una propuesta de dicha explicación.

3. Una explicación funcional ES de la moral humana

Una explicación funcional ES de la moral ha de ser capaz de mostrar que el efecto seleccionado (i. e., la coordinación de un comportamiento cooperativo y prosocial) es adaptativo. Se asume que, en términos generales, eso está demostrado (Hamilton, 1964; Trivers, 1971; Axelrod, 1984; Alexander, 1987). Se argumentará a continuación que existe en el reino animal un rasgo que produce dicho efecto y que, en el caso del ser humano, puede ser considerado como el antecedente evolutivo del juicio moral.

3.1. El juego de las señales y la coordinación de comportamientos cooperativos

Se han desarrollado algunas teorías acerca de cómo la capacidad de emitir señales (*signaling*) influye en las actividades cooperativas en el reino animal (Maynard Smith y Harper, 2003). El mismo Darwin era consciente del poder de las “señales” para convenir acciones que resultaran en una ventaja adaptativa para los individuos involucrados, como puede verse en este ejemplo:

Since monkeys certainly understand much that is said to them by man, and when wild, utter signal-cries of danger to their fellows; and since fowls give distinct warnings for danger on the ground, or in the sky from hawks (both, as well, as third cry, intelligible to dogs), may not some unusually wise ape-like animal have imitated the growl of a beast of prey, and thus told his fellow-monkeys the nature of the expected danger?

This would have been the first step in the formation of a language ([1882] 1989, p. 57).

Este ejemplo muestra que en el reino animal es posible convenir un determinado comportamiento, lo que resulta en una ventaja adaptativa, a través de una señal sin demasiada sofisticación lingüística. Por supuesto que imitar el sonido característico de un predador ya implica cierta complejidad lingüística y cognitiva, pero, si se piensa en el ejemplo más simple de las aves emitiendo sonidos de peligro, entonces es posible admitir que, a través de ciertas señales, en el reino animal es posible convenir comportamientos cooperativos.

La forma más simple de mostrar que pueden evolucionar señales que tengan dicha función es a través del juego de las señales de Lewis (1969). El juego de las señales, en su forma más básica, consiste en dos jugadores, el Remitente (*Sender*) que envía una señal y el Receptor (*Receiver*) que la recibe y actúa. Remitente y Receptor tienen que conspirar para coordinar un acto con un estado del mundo. El acto que tienen que coordinar es aquel que, dado un estado del mundo, resulta mutuamente beneficioso. Solamente el Receptor actúa y solamente el Remitente percibe el estado. Ambos jugadores reciben una ganancia si el Receptor actúa correctamente. En caso contrario ninguno recibe nada. Así, el juego es un juego de pura coordinación, donde los jugadores o tienen éxito o fracasan juntos. Receptor y Remitente no tienen ninguna comprensión previa de qué señal enviada corresponde a qué estado del mundo. De esta manera, el “significado” de las señales es creado por los jugadores cuando el juego alcanza un estado de equilibrio.⁵

El juego involucra dos estados (Estado 1 y Estado 2), dos actos (Acto 1 y Acto 2), y dos señales (designadas como Rojo y Azul para no crear una asociación sugestiva entre señales y estados o actos). Supóngase que el Acto 1 genera beneficio para los dos jugadores en el Estado 1 y supóngase que el Acto 2 genera beneficio para los jugadores en el Estado 2. De esta forma, se tienen cuatro estrategias para el Remitente y cuatro estrategias para el Receptor, y en total 16 combinaciones, tal como muestra la tabla:

⁵ Un grupo de jugadores se encuentra en *equilibrio* si cada uno de los jugadores está tomando la mejor decisión posible, dadas las decisiones de los otros jugadores, siempre y cuando tales decisiones no cambien.

Estrategias del Remitente	Estrategias del Receptor
Rm1) Estado 1: envía 'Rojo'; Estado 2: envía 'Azul'.	Rc1) Realiza Acto 1 si recibe 'Rojo'; Realiza Acto 2 si recibe 'Azul'.
Rm2) Estado 1: envía 'Azul'; Estado 2: envía 'Rojo'.	Rc2) Realiza Acto1 si recibe 'Azul'; Realiza Acto 2 si recibe 'Rojo'.
Rm3) Siempre envía 'Rojo'.	Rc3) Siempre realiza Acto 1.
Rm4) Siempre envía 'Azul'.	Rc4) Siempre realiza Acto 2.

Tabla 1. Estrategias Remitente/Receptor en el juego de las señales.

Las estrategias Rm1 y Rc1 se combinan para generar ganancia para Remitente y Receptor. Las estrategias combinadas Rm2 y Rc2 también son exitosas, pero, a diferencia de la combinación Rm1-Rc1, las señales tienen significados diferentes.

Imáginese ahora una población donde Remitente y Receptor asumen comportamientos condicionales, de tal manera que asumen que, si un individuo es un remitente y percibe determinado estado, envía determinada señal, y si es receptor y percibe determinada señal, realiza determinado acto, etc.

En esta población habrá 16 combinaciones posibles de estrategias Remitente-Receptor (Rm1-Rc1, Rm1-Rc2, Rm1-Rc3, etc.). Y si dicha población consiste en individuos que han adoptado la combinación Rm1-Rc1 (o Rm2-Rc2), en la cual se usa la estrategia Rm1 (o Rm2) cuando se asume el rol de Remitente y Rc1 (o Rc2) cuando se asume el rol de receptor, siempre coordinarán actos que les producen ganancia. Los mutantes que asumen combinaciones diferentes (salvo Rm2-Rc2, como ya se explicó) tendrán menores ganancias, lo que hace que las combinaciones Rm1-Rc1 o Rm2-Rc2 sean las estrategias evolutivamente estables y, por lo tanto, las que muy probablemente evolucionen.

De esta manera, es posible que en el reino animal evolucionen señales que coordinen actos que produzcan beneficios adaptativos para los individuos involucrados. Harms y Skyrms dan algunos ejemplos de

este tipo de señal con contenido semántico primitivo: “The alarm calls of vervet monkeys can be viewed as *indicating* the nature of the predator or as *prescribing* the correct evasive behavior. This sort of signaling system is found in prairie dogs, meerkats, jungle fowl and domestic chickens” (2008, p. 13; énfasis mío).

Se ha mostrado hasta este punto que en el reino animal hay un efecto seleccionado, el cual es el comportamiento cooperativo, y que un rasgo que produce dicho efecto es la capacidad de participar en intercambios coordinados a través de la emisión o recepción de señales.

El argumento que se está desarrollando requiere una aclaración. No se asume que toda señal genera cooperación. Por el contrario, si se admite la naturaleza egoísta de la evolución (Krebs y Dawkins, 1984), la clave para entender la comunicación animal no es la transmisión de información, sino la manipulación. Un animal que emite una señal no busca el beneficio mutuo de emisor y receptor, sino que busca persuadir al receptor para que se comporte del modo más beneficioso para él. Un macho se beneficia si, en vez de usar la fuerza, persuade a una hembra de su especie para que se aparee con él. Las señales son el medio que usa un animal (el emisor) para manipular el comportamiento de otro animal (el receptor).

Sin embargo, esto no significa que no hayan evolucionado señales de cooperación como las analizadas en el juego de las señales. De hecho, las señales que transmiten información falsa o engañosa crean una presión de selección que eventualmente lleva a los receptores a ignorar señales a las que les es perjudicial responder. De esta manera, a medida que una señal engañosa va siendo más frecuente en la población, los receptores van respondiendo cada vez menos a ella. La señal se irá “devaluando” al tiempo que el beneficio que obtienen los emisores mentirosos por emitirla será cada vez menor. En últimas, los receptores dejarán de responder a la señal engañosa. Para que una señal persista en una población, debe ser beneficiosa, en promedio, tanto para el emisor de dicha señal como para el receptor (Greenfield, 2006). Esto significa que, como ya se vio en el juego de las señales de Lewis, la posibilidad de evolución de señales cooperativas está garantizada. Es posible identificar a las señales cooperativas como el rasgo que ha sido seleccionado por el efecto que produce, a saber, coordinar acciones cooperativas. Estas señales adquieren ciertas características en el ser humano, dando paso a la moral. En lo que sigue se intenta mostrar cómo las señales cooperativas se vinculan con la moral.

3.2. La señal cooperativa y la fuerza práctica del juicio moral

No toda señal cooperativa está vinculada con el juicio moral o constituye un lenguaje protomoral. Las abejas y las hormigas, por ejemplo, exhiben comportamientos cooperativos mediados por señales, pero difícilmente se dice que son animales morales o que poseen lenguaje protomoral.

Lo que le da el carácter de moral a las señales cooperativas en el ser humano y de protomoral en los primates superiores no humanos es la forma particular en que funcionan tales señales en cada especie. Y esto depende de la historia evolutiva particular de la especie. Esta es la razón por la cual uno de los objetivos de este trabajo es defender la idea de que la explicación funcional evolutiva de la moral no solamente ha de encontrar el rasgo que produce el efecto cooperativo (explicación funcional del efecto seleccionado), sino que también ha de dar cuenta de la forma particular en que se produce el efecto seleccionado (explicación funcional del rol causal). Para lograr esto, es necesario evaluar la forma particular en que funciona la señal cooperativa tanto en primates superiores no humanos como en seres humanos.

Para empezar, puede decirse que, dada la selección de una estrategia Remitente-Receptor que aumenta la aptitud biológica de ambos, cuando la estrategia es violada, es decir, cuando el Receptor no realiza la acción esperada una vez emitida determinada señal, se activa una alarma en el emisor —tanto en primates superiores no humanos como en seres humanos— que le indica que debe actuar para que se cumpla la respuesta esperada.

Esto muestra que la señal cooperativa posee dos características importantes: por una parte, posee condiciones objetivas de verdad, pues señala que hay un estado objetivo del mundo (Estado 1) que la señal representa, y, en segundo lugar, guía directamente la acción, pues hay una acción (Acción 1) que es adaptativa frente al estado del mundo que la señal indica.

Luego, la señal cooperativa posee un carácter indicativo y prescriptivo. Este carácter indicativo y prescriptivo podría ser el origen del carácter autoritativo del juicio moral, pues el hecho de que la señal indique un estado convencionalmente objetivo del mundo hace que la señal tenga el carácter de inevitable. No depende de mis intereses particulares aceptar o no el estado de cosas que indica la señal cooperativa. Y, sumado a esto,

dado que este estado de cosas prescribe un curso de acción determinado, la señal cooperativa posee también autoridad. Recuérdese que, para Joyce, estos son los elementos esenciales del juicio moral.

Podemos reemplazar las señales del juego de las señales con juicios morales tales como “robar es malo”, y se siguen manifestando las dos características importantes mencionadas que revelan la naturaleza del juicio o cognición moral: en primer lugar, las señales poseen condiciones objetivas de verdad (hay un “Estado 1” objetivo del mundo que la señal representa) y, en segundo lugar, guían directamente la acción, pues hay una Acción 1 que es adaptativa frente al Estado 1 (Harms y Skyrms, 2008, p. 12).

Sin embargo, en términos generales, esta característica se le puede atribuir de manera abstracta a toda relación de comunicación en la que una señal lleve a determinado curso de acción. Por esta razón resulta poco convincente admitir que esta es la esencia del lenguaje moral (o protomoral).⁶ Pero si este tipo de señal sirve como medio de expresión de ciertas tendencias evaluativas básicas, a la vez que fortalece los lazos cooperativos entre los individuos de una especie, entonces empieza a adquirir la forma de norma o imperativo moral, con su respectivo carácter autoritativo. La fuente de este carácter autoritativo parece haberse *interiorizado* en el caso del ser humano. Ya se dijo que, frente a una respuesta incorrecta en la emisión de una señal cooperativa, se activa una alarma que indica que hay que corregir la respuesta. En los primates no humanos, el garante de que se corrija la respuesta suele ser el macho dominante o la comunidad o los individuos adultos, en general. En los seres humanos, la comunidad y la autoridad externa juegan un papel importante, pero, gracias a la capacidad de lenguaje y memoria, el ser humano puede “llevar consigo” la señal cooperativa, por decirlo de alguna forma, prescindiendo de la autoridad externa. Esto hace también que el ser humano pueda ser al mismo tiempo emisor y receptor de la señal cooperativa, pues es capaz de entablar una especie de diálogo interno consigo mismo. Igualmente, como se verá, el ser

⁶ Si bien el carácter indicativo-prescriptivo de la señal cooperativa es una característica que, en términos generales, comparten todas las formas de comunicación a través de señales, no sería extraño que, tratándose de algo tan serio como la cooperación, la evolución hubiera cooptado esta característica para generar en la psique humana la percepción de las normas y juicios morales como hechos objetivos del mundo (cfr. Ruse y Wilson, 1996; Joyce, 2006; Kitcher, 2006).

humano es capaz de dar expresión de diversas formas a sus tendencias evaluativas básicas, a través de señales cooperativas, generando en todo caso cooperación, lo que da como resultado sistemas de señales distintos o, para ponerlo en términos morales, generando pluralismo moral.

En suma, las señales cooperativas poseen un carácter indicativo-prescriptivo que, acompañado de ciertas tendencias evaluativas básicas, puede explicar la cognición moral. Por ahora es necesario revisar de qué manera las tendencias evaluativas básicas de los primates superiores no humanos influyen en el funcionamiento de las señales cooperativas, y hasta qué punto las determinan. Esto servirá para preparar el camino de la explicación funcional del juicio moral humano.

La “fuerza práctica” que Joyce atribuye a los juicios morales se puede explicar, en principio, de la siguiente manera: el carácter inescapable del juicio moral posee su equivalente en las señales morales en el hecho de que tales señales poseen condiciones objetivas (aunque convencionales) de verdad. Por otra parte, toda norma o juicio moral expresa o indica una situación que, de darse o presentarse, prescribe una acción determinada; justamente este carácter prescriptivo es a lo que Joyce se refiere cuando señala que los juicios morales son “autoritativos”, pues mandan o prescriben ciertos comportamientos o actitudes. Ahora se pasará a investigar la forma en que funciona la señal moral en los primates superiores no humanos. Si es cierto que el comportamiento social de los chimpancés posee los bloques fundamentales (*building blocks*) del comportamiento moral humano (cfr. De Waal, 2007), está justificado investigar la forma en que funciona la señal moral en los primates superiores no humanos. Esto permitirá determinar la teoría funcional biológica adecuada para dar cuenta de la moral humana.

3.3. La señal cooperativa en primates superiores no humanos

Hay evidencia de que algunos primates vocalizan de una forma característica ciertos sonidos cuando son víctimas de agresiones por parte de individuos de su misma especie. En los chimpancés, la estructura acústica de los gritos de la víctima dice algo sobre la severidad de la agresión, pero también indica la composición de la audiencia: los gritos indican una agresión más severa cuando la audiencia está compuesta de individuos de alto rango, con independencia de la naturaleza física del ataque (Slocombe y Zuberbühler, 2007).

Junto con los chimpancés, otros primates vocalizan, no solamente con la intención de amedrentar a su oponente, sino también para

recibir apoyo de la audiencia, en el sentido de que buscan que algunos individuos se adhieran a su causa o que otros de mayor rango cumplan funciones policivas, o también para recibir consuelo (Von Rohr, Burkart, y Van Schaik, 2011).

Existe evidencia también de un tipo de comportamiento exhibido por el mono capuchino marrón (*cebus apella*) que cuenta como señal cooperativa. Brosnan y De Waal (2003) llevaron a cabo un experimento en el cual los experimentadores, luego de hacer que los monos realizaran cierta tarea que exigía igual esfuerzo por parte de los individuos involucrados, les ofrecían una recompensa. En un caso, la recompensa ofrecida era igual para los implicados en la tarea, mientras que, en otra situación, frente al mismo esfuerzo realizado, les ofrecían recompensas desiguales, ofreciéndole a algunos monos tajadas de pepino y uvas a otros, las cuales son muy codiciadas entre ellos. El resultado del experimento fue que, frente a recompensas inequitativas, los monos rechazaban la recompensa, mientras que, cuando la recompensa era equitativa, sin importar qué tan grande o apetecida fuera, los individuos la aceptaban sin vacilar. Esto sugiere, según los investigadores, que existe cierto sentido de la equidad en la distribución de los recursos entre los monos capuchinos marrón. Y sugiere que este sentido de la equidad es relativo respecto a la recompensa obtenida por sus compañeros. Así, frente a un estado de inequidad en la distribución, los monos emiten una señal —manifestada como un comportamiento de protesta—, que ha de ser interpretada por el experimentador (en el juego de las señales, el “Receptor”) como “llevar a cabo la distribución equitativa de la recompensa”.

Nótese que en estos ejemplos existen unos requisitos para que las señales cooperativas resulten efectivas y produzcan la coordinación propia del comportamiento cooperativo. Estos requisitos tienen que ver con ciertas características propias de la historia evolutiva contingente de los primates. Así, por ejemplo, en el caso de la aversión a la inequidad, existe la capacidad de medir la justicia de la recompensa, no en relación con el esfuerzo realizado, sino más bien en relación con lo que recibe el compañero de trabajo.

Y, respecto al primer ejemplo, en el que la señal moral se vocaliza, se evidencia que existe en los individuos de menor rango la tendencia a refrenar sus impulsos en presencia de una autoridad. Así, por ejemplo, en macacos y otros primates, los machos de menor rango varían su comportamiento dependiendo de la ausencia o presencia del macho

alfa. Tan pronto como el macho alfa da la espalda o se retira, los otros machos se aproximan a las hembras. Cuando el macho alfa está ausente, algunos machos de menor rango se atreven incluso a realizar el gesto de rebote, típico del macho dominante. Luego de estos episodios, los investigadores han podido constatar que, al volver a reunirse con el macho alfa, los suplantadores se muestran excesivamente nerviosos, lo que se evidencia en una exagerada sumisión a través de un gesto exagerado de exhibición de sus dientes. Algunos investigadores han interpretado este comportamiento como el reconocimiento implícito por parte de los individuos de menor rango de que han violado una norma de conducta (Coe y Rosenblum, 1984).

Así como existe la tendencia a refrenar sus impulsos en presencia de una autoridad, parece existir también evidencia de que la autoridad en los primates superiores no humanos se manifiesta de manera imparcial.⁷ Por ejemplo, los chimpancés machos de alto rango frecuentemente ejercen un rol de control al intervenir en peleas entre hembras o machos más jóvenes del grupo (De Waal, 1982). Y, a diferencia de la mayoría de los individuos que apoyan a sus parientes o amigos en este tipo de peleas, existe amplia evidencia de que los machos que ejercen el rol de control ignoran sus vínculos sociales al intervenir en las peleas. Esta forma de control imparcial se ejerce también en los chimpancés que se encuentran en estado salvaje (Von Rohr, Burkart, y Van Schaik, 2011).

En suma, estos estudios ponen en evidencia que la señal cooperativa en los primates superiores no humanos está determinada por ciertas actitudes y tendencias conductuales, que son las que determinan la forma en que la señal cumple su función.

Para que la señal funcione, deben existir ciertos prerequisites que tienen que ver con la disposición natural de los individuos a comportarse de maneras determinadas. En los ejemplos recién ofrecidos se vio que existe una tendencia a refrenar los impulsos en presencia de la autoridad; la autoridad tiende a actuar de manera imparcial, y también se vio que el castigo es el mecanismo usado para enseñar que existe una

⁷ Es necesario ser cuidadosos con la forma en que se interpretan estas manifestaciones de “imparcialidad” de la autoridad. Tal como Rosas lo señala, la imparcialidad no necesariamente tiene que ser vista como una tendencia a administrar justicia, sino que podría interpretarse también, entre otras cosas, como la acción del macho alfa destinada a evitar que en los conflictos haya un ganador y, a la postre, un contendiente a su autoridad (cfr. 2005, pp. 98-99).

normatividad social y que su incumplimiento lleva a la intervención de la autoridad.

4. La explicación funcional plural de la moral

Existe un rasgo (la señal cooperativa) que produce un efecto, a saber, el de coordinar comportamientos prosociales y cooperativos, y es por este efecto que dicho rasgo ha sido seleccionado. Esto significa que el efecto mencionado (explicación ES) es la función que se le atribuye a la señal cooperativa. Pero, en la medida en que el funcionamiento de la señal cooperativa está condicionado por unas tendencias evaluativas básicas, contingentes a la especie,⁸ como se vio en el caso de los primates superiores no humanos, una explicación funcional de la forma en que funciona la señal cooperativa necesita dar cuenta de tales tendencias evaluativas o de la forma en que estas moldean la emisión y aceptación de la señal.

Esta conclusión, que se obtiene al evaluar la forma en que funciona la señal cooperativa en los primates no humanos, aplica igualmente para el hombre, asumiendo que hay motivos para creer que tales primates revelan aspectos ancestrales de la proto-moralidad humana (De Waal, 2007).⁹ La explicación funcional del efecto seleccionado (ES), en la medida en que solamente da cuenta del efecto por el cual el rasgo al que le atribuimos la función de la moral ha sido seleccionado, no es capaz de dar cuenta de la forma en que se relaciona tal efecto con otros aspectos de la psicología humana. Estos aspectos afectan el funcionamiento del efecto seleccionado. Por tal motivo, es necesario recurrir a una explicación funcional alternativa.

Al principio de este artículo se dijo que en Biología son necesarias explicaciones funcionales que combinan explicaciones ES con explicaciones RC (Amundson y Lauder, 1994; Huneman, 2013; Jackson

⁸ Tales tendencias son el resultado de una historia evolutiva particular y contingente, de manera que no pueden trasladarse de una especie a otra (cfr. Ruse y Wilson, 1986, p. 173 y Street, 2006, p. 117).

⁹ De Waal hace esta afirmación basándose en manifestaciones de reciprocidad, comportamiento de consuelo, aversión a la inequidad y empatía, entre otras, que se encuentran presentes y documentadas en los chimpancés. A pesar de esto, De Waal admite que los chimpancés no hacen juicios morales (cfr. 1996, p. 209).

y Fry, 2016). Este pluralismo funcional es el que resulta adecuado para explicar la función del juicio moral.

Una explicación ES del corazón explicará por qué ha sido seleccionado el corazón (a saber, por su efecto de hacer circular la sangre). Pero el efecto funcional del corazón, el cual es hacer circular la sangre, necesita, para su explicación, determinar la forma en que tal efecto se integra con otros sistemas, como el muscular o el respiratorio, por ejemplo. De otra forma, no se podrá entender por qué, entre otras cosas, el corazón bombea sangre más rápido cuando los músculos están en movimiento, o por qué el corazón posee un sistema exclusivo que lo conecta directamente con los pulmones. De manera similar, en el caso de la explicación funcional del juicio moral, dicha explicación deberá tener en cuenta la interacción del efecto seleccionado del juicio moral con las tendencias evaluativas humanas básicas.

La concepción funcional plural de la moral humana está implícita en las teorías de algunos filósofos. Para Gibbard, por ejemplo, la explicación funcional ES no es suficiente, pues la simple tendencia a expresar normas no basta para explicar la moral, sino que es necesario tener en cuenta que los juicios normativos morales responden a deseos y ventajas, las cuales han de ser adaptativas (cfr. 1990, p. 312). Corresponde entonces a una explicación funcional RC caracterizar de la manera más precisa posible cuáles son esos deseos y tendencias particulares que determinan o moldean la capacidad de coordinar comportamientos cooperativos.

Igualmente, para Wong (2006, pp. 39-40) la moral busca promover y sustentar la cooperación social y el florecimiento individual (explicación funcional ES), pero existen necesidades físicas, sociales y de conocimiento, entre otras, que han de satisfacerse si la moral ha de cumplir adecuadamente su función (explicación funcional RC). Esta concepción de la moral hace posible hablar del cumplimiento o no de la función de la moral, y establece que, para hacerlo, es necesario identificar algunas de las mencionadas propensiones evaluativas básicas del ser humano. En lo que sigue se intenta esbozar una posible explicación funcional RC del juicio moral humano.

4.1. La explicación RC del juicio moral

Las tendencias evaluativas básicas del ser humano moldean el efecto producido por la señal cooperativa y determinan si el efecto producido por dicha señal funciona o no, es decir, si promueve el efecto cooperativo esperado.

La *tendencia evaluativa básica* es diferente al juicio evaluativo. Los juicios evaluativos son un conjunto más amplio de juicios que incluye las normas o juicios morales, las cuales poseen elementos evolutivos ancestrales, pero también elementos racionales y culturales. La tendencia evaluativa básica puede ser definida como una “proto”-forma de juicio evaluativo que consiste en un impulso —no reflexivo ni lingüístico— a llevar a cabo cierto comportamiento que es demandado o necesario. Este carácter demandante está dado porque las tendencias evaluativas solucionan determinados problemas evolutivos, aumentando la aptitud biológica de los individuos.

En el caso de los primates superiores no humanos se vio que, aunque no hacen juicios evaluativos, sí poseen ciertas predisposiciones o impulsos, como el de estar dispuestos a someterse a cierta autoridad, el de ser benevolentes con sus crías o el de protestar frente a una repartición inequitativa de la recompensa, entre otros.

Se dijo que la explicación RC del juicio moral ha de enfocarse en determinar la forma en que el efecto cooperativo y prosocial del juicio moral es moldeado o influenciado por las tendencias evaluativas básicas del ser humano, de manera que la función del juicio moral está determinada y moldeada por dichas tendencias.

Para determinar cuáles son algunas de tales tendencias en el caso del ser humano, resulta de utilidad remitirse al estudio meta-empírico llevado a cabo por Haidt y Joseph (2004), en el cual estos investigadores, luego de consultar algunos trabajos que describen lo que es universal en la moral humana, otros pocos que se centran en lo que es diferente y un trabajo que investiga los elementos básicos de la moral que compartimos humanos y primates no humanos, llegan a la conclusión de que existen cuatro módulos¹⁰ o dominios morales básicos en el ser humano. Según estos autores, los seres humanos:

¹⁰ Se entiende por “módulo” un sistema de procesamiento diseñado para lidiar con problemas que se les presentaban a las especies en su ambiente ancestral por muchas generaciones. Los módulos consisten en sistemas de programación automática que, frente a determinado estímulo (*input*), producen automáticamente una respuesta (*output*). La ventaja de estos módulos consiste en producir respuestas rápidas e intuitivas ante los problemas que requerían solución inmediata. El primero en proponer la modularidad respecto a procesos perceptuales fue Jerry Fodor (1983).

[...] come equipped with an intuitive ethics, an innate preparedness to feel flashes of approval or disapproval toward certain patterns of events involving other human beings. The four patterns for which we believe the evidence is best are those surrounding suffering, hierarchy, reciprocity, and purity (2004, p. 58).

Estos cuatro módulos no son exhaustivos del pensamiento moral, pero cada uno es adaptativo. Así, por ejemplo, el módulo que detecta el sufrimiento de los demás jugó un papel fundamental para permitir que las madres detectaran que sus hijos estaban bajo condiciones amenazantes y de estrés. Madres con una capacidad de detectar más fácilmente el sufrimiento en sus hijos seguramente tendrían mayores probabilidades de tener descendencia numerosa. Con el tiempo se desarrolló un sistema de comunicación en el cual sofisticadas señales de estrés en los hijos disparaban automáticamente la ayuda materna.

Respecto al rasgo psicológico de categorizar las relaciones humanas en términos jerárquicos, es evidente que este rasgo ayuda a los animales a vivir en grupos sociales y facilita la posibilidad de dominar a otros y de ser aceptados como parte del grupo. Dada una distribución inequitativa de fortaleza, habilidades y suerte, los individuos que tenían las reacciones emocionales correctas lograron desenvolverse de manera exitosa en medio de este juego desigual de fuerzas, a diferencia de los individuos que rehusaban subordinarse o que no mostraban las reacciones apropiadas en medio de situaciones en las que estaba en juego la dominación.

Respecto al carácter adaptativo de la reciprocidad ya se dijo lo necesario. Pero, además de estos tres dominios o módulos morales, Haidt y Joseph postulan un cuarto módulo, el cual consiste en la moralización del cuerpo y de las actividades corporales, tales como la menstruación, el comer, la limpieza corporal, el sexo y la manipulación de cadáveres, entre otros (cfr. 2004, p. 60). Este dominio moral está relacionado explícitamente con la regulación de la pureza y de la polución. Esta regulación posee claramente la ventaja adaptativa de lidiar de manera ventajosa con gérmenes, microbios y parásitos potencialmente peligrosos para el individuo.

Se han mencionado algunas de las tendencias evaluativas básicas del ser humano. Sin embargo, la explicación funcional RC explica cómo el efecto seleccionado del juicio moral (el efecto cooperativo y prosocial)

se relaciona con estas tendencias básicas, o, dicho de otra forma, la explicación funcional RC explica de qué forma estas tendencias influyen en el éxito o fracaso del juicio moral para lograr un comportamiento cooperativo y prosocial.

Lo interesante de caracterizar las tendencias evaluativas básicas en el ser humano no consiste simplemente en explicar por qué son adaptativas, sino más bien en exponer cómo se expresan a través de ciertos mecanismos automáticos, pues son justamente estas particularidades las que determinan el funcionamiento del juicio moral.

4.1.1. La expresión de las tendencias evaluativas básicas

En el caso del sufrimiento, existe un mecanismo, descubierto por el equipo del neurobiólogo Rizzolatti (2004), que no solamente permite explicar los mecanismos neuronales que actúan en las acciones empáticas que involucran la percepción de que otros están sufriendo, sino que en general facilita la comprensión de acciones básicas de los demás seres humanos. Según dicho equipo, existen unas neuronas “espejo” cuya función parece consistir en reflejar la actividad que se está observando. Estas células nerviosas se activan cuando otra persona sufre, cuando se ejecuta determinada acción o cuando se ve que otra persona la realiza. Tales neuronas logran que se tenga una comprensión inmediata e intuitiva de algunas acciones básicas de los demás. El mecanismo es útil para adjudicar intenciones a los demás, similares a las que posee quien las adjudica.

Hay varias conductas automáticas que expresan la tendencia humana a jerarquizar. Así, por ejemplo, de manera automática, cuando se forman vínculos sociales, el ser humano se adhiere a un colectivo, lo que inmediatamente le permite crear bandos y distinguir entre el “nosotros” y el “ellos” (Tahjfel, 1981). A su vez, la jerarquización que hace el ser humano sobre el mundo social depende de la imagen que crea de sí mismo. Es a partir de un autorrelato, el cual responde a la pregunta sobre cuál es su principal rol en el mundo, que el ser humano mide la jerarquía social de los que se relacionan con él. Esta autoimagen provee la base de la autoestima, importante para mantener relaciones sociales estables (Rogers, 1961).

El sentido humano de reciprocidad está fuertemente influenciado por la tendencia humana a valorar la recompensa en términos relativos. Existe un juego llamado Ultimátum (Camerer y Thaler, 1995) en el cual a un jugador (oferente) se le propone que reparta una determinada

cantidad de dinero con otro jugador (respondedor) según le convenga, haciendo una única y definitiva propuesta. El respondedor podrá aceptar o no dicha propuesta. En caso de no aceptar, ningún jugador gana nada. Pero si acepta, se reparte la cantidad según la propuesta realizada. El experimento muestra que los criterios de elección se basan en criterios de justicia y equidad que no tienen que ver con el beneficio, pues existe la tendencia a realizar propuestas justas y a rechazar las propuestas en las que el oferente propone una ganancia demasiado alta para él en comparación con la cantidad que le ofrece al respondedor. La ganancia y lo que se considera una repartición equitativa no se mide en términos absolutos, sino más bien en relación con lo que obtienen los demás.

Igualmente, la tendencia humana a regular las relaciones sociales a través de la reciprocidad está ampliamente influenciada por la propensión humana a crear relaciones estables y duraderas (Schacter, 1959; Maslow, 1968; Baumeister y Leary, 1995), de tal manera que el ser humano está dispuesto a adaptar sus creencias, actitudes y comportamientos a los del grupo al cual pertenece, pues, de no hacerlo, teme ser excluido del grupo. De esta manera, el criterio de reciprocidad está determinado por la influencia del grupo, mereciendo premio y alabanza aquello que aprueba el grupo y castigo o reproche lo que desaprueba.

Estos son apenas unos pocos ejemplos de la forma en que se expresan las tendencias evaluativas básicas humanas por medio de mecanismos automáticos. Ahora es necesario ver cómo esto influye en la determinación de la moral humana.

4.1.2. La explicación funcional RC y el deber moral

En los primates superiores no humanos existe una propensión a obedecer a la autoridad en la forma del macho dominante o de un adulto que interviene para restablecer la armonía social y coordinar actos cooperativos. Sin embargo, en el ser humano esta autoridad se ha “interiorizado”, de manera que percibe al mandato moral como proveniente de su propio interior.

¿Cómo se explica esta interiorización de la autoridad en el ser humano, de manera tal que la norma se muestra obligatoria, no por temor al castigo de una autoridad externa, sino como si algo del interior moviera al sujeto a cumplir la norma moral? La autoridad moral puede ser explicada no solamente por el carácter indicativo y prescriptivo de la señal cooperativa, sino principalmente por la forma particular en que se expresan las tendencias evaluativas básicas de jerarquización

y reciprocidad en el ser humano. Se dijo que la reciprocidad está fuertemente influenciada por la tendencia a pertenecer a un grupo. Esta tendencia a la “filiación” se manifiesta de manera particular en los niños, lo cual está documentado por varios estudios de la psicología del desarrollo. Así, por ejemplo, los niños dicen “mentiras inocentes” (*white lies*) si eso les permite entablar relaciones más duraderas (cfr. Over, 2016, p. 3).

También hay experimentos que muestran cómo los niños, una vez que han sido incluidos en un grupo, buscan adaptar su comportamiento al juicio los de los demás. Por ejemplo, Haun y Tomasello (2011) mostraron que niños de cuatro años a los cuales se les presentó una versión modificada de la tarea de Asch¹¹ se conformaban a la opinión incorrecta de los miembros de su grupo aproximadamente un tercio de las veces, y tres cuartos de los niños lo hacían por lo menos en una de las pruebas. En un segundo experimento, a los niños que participaron en la tarea de Asch se les permitió dar sus reportes en privado, y los niveles de conformidad con el grupo fueron mucho menores. Una posible interpretación de los resultados es que los niños *buscan ser aceptados* por el grupo y evitan su desaprobación.

Existe en el ser humano la necesidad de aceptación, de manera que, por temor al rechazo, o mejor, por la necesidad de ser aceptado, interioriza las normas morales que la educación y la formación le han impuesto desde niño. Si se transgreden tales normas, surge la culpa. La culpa puede ser definida como el juicio de que se merece castigo por transgredir una norma (Joyce, 2006). Tal como decía Darwin, el individuo imagina que, si los demás descubren que ha transgredido la norma, será

¹¹ Salomón Asch (1951) diseñó un experimento en el que solicitaba a un grupo de sujetos que indicase, respecto de un conjunto de tres líneas de diferente tamaño, cuál de ellas se asemejaba más a una línea estándar o de prueba. Los sujetos experimentales no siempre decían la respuesta correcta. El experimento trataba de ver cómo la presión de grupo fuerza a variar los juicios. Las pruebas del experimento se realizaban a un grupo de unas seis u ocho personas, de las cuales solo uno era verdaderamente un sujeto experimental, ya que los demás (sin saberlo el sujeto experimental) eran cómplices del investigador. Durante algunos de los ensayos de las pruebas (ensayos críticos), los cómplices daban respuestas claramente erróneas. Muchos de los sujetos experimentales optaron por decir lo mismo que los cómplices del experimentador, es decir, optaron por las respuestas falsas; estuvieron de acuerdo con la respuesta equivocada el 37% de las veces.

excluido del grupo. Pero aun si no es descubierto, el individuo siente la necesidad de corregir sus acciones, reparar y conformarse a los preceptos del grupo, pues solo así da expresión a esa propensión psicológica de sentir que posee una “filiación” con su grupo.

Pero el deber o autoridad moral no solamente se explica en función de la tendencia del ser humano a la reciprocidad. Algunos psicólogos han percibido que ciertas personas responden defensivamente a fallas morales, mientras que otras responden queriendo enmendar sus fallas (Gausel y Leach, 2011). La forma en que tradicionalmente se explica esta diferencia es argumentando que, cuando la falla moral lleva a la culpa, generalmente hay una actitud de reparación, mientras que, si la falla lleva a la vergüenza, se activan mecanismos de defensa. La razón por la que se activan mecanismos de defensa es que, en el caso de la vergüenza, el individuo transgresor juzga que los demás han detectado un defecto en su propio ser.

Ya se señaló que parte fundamental de la tendencia a jerarquizar en el ser humano depende de su autoimagen, entendida como la concepción de su propia identidad. La emoción moral de la vergüenza parece surgir cuando tal identidad se ve amenazada. Esta autoimagen provee la base de la autoestima, la cual es importante para mantener relaciones sociales estables (Rogers, 1961). Por lo tanto, en el caso de la vergüenza está en juego la autoimagen. El ser humano tiende a construir un relato acerca de sí mismo, de quién es (a crear un lugar en la jerarquía social), de manera que las complejas explicaciones que dan sentido a su mundo incluyen expectativas respecto a cómo debe comportarse, generando la necesidad de mantener una imagen propia en sus acciones recíprocas para así mantener su lugar o jerarquía frente a los demás.

Es así como la auto imagen y las expectativas que el ser humano crea sobre sí mismo se convierten en estándares que debe cumplir, en normas que poseen autoridad sobre sus actos. De esta manera, tanto la tendencia a reciprocitar como la propensión a jerarquizar pueden dar cuenta de la autoridad moral.

Pero la expresión de las tendencias evaluativas básicas no solamente permite explicar el deber moral, sino que, al ser dichas tendencias expresadas en contextos culturales contingentes distintos, permite explicar cómo es que juicios morales incompatibles pueden cumplir con la función del juicio moral. Para mostrar esto, es necesario analizar cómo las conductas automáticas que expresan las tendencias evaluativas, al ser influenciadas por factores socioculturales o económicos

contingentes, pueden producir sistemas morales diferentes o incluso inconmensurables.

4.2. La explicación funcional plural de la moral y la ambivalencia moral

Una explicación funcional evolutiva como la ofrecida en este artículo debe ser capaz de explicar la *ambivalencia moral*, entendida como el fenómeno según el cual un individuo llega a entender el punto de vista moral contrario al que sostiene, al punto de que su creencia absoluta en su punto de vista moral llega a verse desestabilizado (cfr. Wong, 2006, p.5). Un buen ejemplo de este tipo de desestabilización es el que da Sartre (1946, pp. 43-45). Un pupilo del filósofo tiene la obligación, por una parte, de acompañar a su anciana madre, la cual requiere para su felicidad que su hijo esté cerca de ella. Pero, por otra parte, el pupilo tiene el deber, como ciudadano, de enrolarse en el ejército libertador francés en Inglaterra. ¿Cuál es la decisión moralmente correcta? Existen poderosas razones para justificar una u otra decisión, de suerte que la creencia absoluta en un punto de vista moral se ve desestabilizado.

La concepción evolutiva funcional del juicio moral establecida en este artículo explica este fenómeno argumentando que es posible que se de expresión a determinadas tendencias evaluativas básicas en contextos distintos, generando juicios morales incompatibles, pero cumpliendo aun así con la función de coordinar comportamientos prosociales y cooperativos. Esto permite ver más claramente la diferencia entre tendencia evaluativa básica y señal cooperativa. Podría argumentarse que las tendencias evaluativas, al pertenecer al dominio de la moral, generan por sí solos comportamientos cooperativos, lo que haría innecesaria la explicación funcional ES de la señal cooperativa. Sin embargo, la sola expresión de las tendencias evaluativas no genera cooperación en todos los contextos. Dependiendo del contexto, pueden evolucionar señales cooperativas que dan expresión de manera diferente a la misma tendencia evaluativa básica.

En el ejemplo de Sartre se encuentra en juego la la tendencia humana a jerarquizar a través de la adhesión a un colectivo, lo que inmediatamente le permite al ser humano crear bandos y distinguir entre el “nosotros” y el “ellos” (Tahjfel, 1981). La solución al dilema moral del pupilo de Sartre dependerá de cómo las condiciones contextuales hayan permitido incluir en sus preocupaciones morales a los demás, de

manera que considere que sus compatriotas franceses forman parte del “nosotros” tanto como su propia familia.

A pesar de que existe esa fuerte tendencia a defender los intereses de quienes están incluidos en el conjunto del “nosotros”, la distinción misma está claramente influenciada por contextos y procesos culturales contingentes, como por ejemplo el proceso globalizador. Ya se dijo que, en el ser humano, el sentido de la equidad es relativo a la recompensa que reciben los demás (Camerer y Thaler, 1995). La globalización, por ejemplo, ha ampliado el rango de seres humanos con los cuales es posible medir la equidad. Las encuestas revelaron que, luego de la masificación de la televisión, la gente empezó a dar mayor importancia a la comparación de su calidad de vida con ciudadanos de otros países y no con ciudadanos del mismo país (cfr. Friedman, 2010, p. 347). Esto quiere decir que el rango del “nosotros” se amplía considerablemente gracias a la globalización. Pero, en la medida en que tal ampliación está condicionada por procesos contingentes,¹² es de esperarse entonces que el ser humano pueda coordinar comportamientos cooperativos y prosociales en contextos no globalizados, lo cual significa que, dependiendo del contexto, se pueden producir juicios morales distintos, pero igualmente eficaces a la hora de cumplir su función.

Otra forma de mostrar la contingencia de la determinación consiste en argumentar que muchas veces no existe una ampliación del rango de lo que se considera el “nosotros”, sino que hay un desplazamiento de las intuiciones morales ancestrales. Así, por ejemplo, a quien piensa que no matar brujas hoy en día es una muestra de que el ser humano ha ampliado el círculo de las preocupaciones morales podría replicársele que no hay tal ampliación, sino que lo que hay es un conocimiento de un hecho del mundo: no existen mujeres que han adquirido poderes malignos sobrenaturales otorgados por el diablo. Sin embargo, si fuera un hecho comprobado que tales mujeres existieran, seguramente muchos creerían justificado eliminar a tales seres. No se trata entonces de una ampliación del círculo, sino más bien de un desplazamiento. La —muchas veces— falsa atribución de la mayoría de las conductas delictivas de un país a los inmigrantes justifica la xenofobia de manera

¹² James argumenta (2002), a través de ejemplos históricos, que la globalización no es inevitable y que nada impide que ocurra una crisis económica a escala mundial (como la ocurrida durante La Gran Depresión en 1929) que la ponga en riesgo.

similar a como alguna vez una falsa creencia justificó el temor a las brujas. La tendencia evaluativa es básicamente la misma y lo que cambia es su contenido, determinado por el contexto. El contexto varía, pero las tendencias evaluativas ancestrales son las mismas. Y, si —tal como se ha asumido en la explicación funcional de la moral ofrecida en este artículo— el funcionamiento del juicio moral está determinado principalmente por la expresión de las tendencias evaluativas básicas del ser humano, el contexto puede variar y producir juicios morales diferentes pero que cumplen su función de manera adecuada.

5. Conclusiones

La teoría funcional que mejor se adapta a la concepción de la moral como un rasgo o rasgos que han evolucionado con un propósito es la teoría del *efecto seleccionado* (ES). Sin embargo, se vio que adoptar la teoría funcional ES ha de implicar poder dar cuenta de cuál es el rasgo o conjunto de rasgos que producen el efecto de coordinar acciones cooperativas y prosociales, y ha de poder explicar las características que se atribuyen al pensamiento moral, sintetizadas en la forma del juicio moral, a saber, en su autoridad moral o fuerza práctica.

Se postuló a la señal cooperativa como ese rasgo que coordina acciones a través de un elemento indicativo (lo cual la hace parecer inevitable) y un elemento prescriptivo (que explica su apariencia de autoritativa).

Al evaluar la forma en que funciona la señal cooperativa en los primates superiores no humanos, se puso en evidencia que la explicación funcional ES no es suficiente para dar cuenta de la moral, sino que, en la medida en que la señal moral funciona moldeada por las particularidades de las tendencias evaluativas de los individuos de la especie, entonces se requiere también de una explicación funcional del *rol causal* (RC), la cual consiste básicamente en una explicación de la forma en que el efecto seleccionado de la señal cooperativa (i. e. el comportamiento coordinado cooperativo) interactúa con las tendencias evaluativas básicas para cumplir su función. Por lo tanto, se identificaron algunas de tales tendencias evaluativas básicas en el ser humano y se exploraron algunos de los mecanismos que las expresan para mostrar de qué manera estos influyen en la determinación de la moral. Esta exploración permitió caracterizar de mejor forma la autoridad moral en el ser humano, pues la forma particular en que se manifiesta la tendencia a crear relaciones recíprocas y a jerarquizar permite dar cuenta de

emociones morales como la culpa y la vergüenza, fundamentales para explicar la autoridad moral. Tal explicación RC del juicio moral permitió también dar cuenta del fenómeno de la ambivalencia moral y darle sentido a la idea de que, en el caso del juicio moral, este puede cumplir su función de maneras diferentes, siempre limitado por tener que dar expresión a las tendencias evaluativas básicas. Las diferentes maneras en que se cumple la función del juicio moral están determinadas por elementos culturales diversos, los cuales pueden generar juicios morales incompatibles pero igualmente exitosos a la hora de cumplir su función dependiendo del contexto. Una explicación funcional de la moral ha de ser plural, en el sentido de que la explicación RC complementa a la explicación ES.

Lo que permanece aún sin explorar y ha sido apenas mencionado en este trabajo es el rol que juegan la razón y la emoción en la determinación de la moral humana.

Bibliografía

- Alexander, R. D. (1987). *The Biology of Moral Systems*. Aldine de Gruyter.
- Amundson, R. y Lauder, G. V. (1994). Function Without Purpose: The Uses of Causal Role Function in Evolutionary Biology. *Biology and Philosophy*, 9, 443-469.
- Asch, S. (1951). Effects of Group Pressure on the Modification and Distortion of Judgements. En H. Guetzkow (ed.), *Group, Leadership and Men*. (pp. 177-190). Carnegie Press.
- Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.
- Ayala, F. J. (2010). The Difference of Being Human: Morality. *Proceedings of the National Academy of Sciences*, 107, 9015-9022.
- Baumeister, R. F. y Leary, M. R. (1995). The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. *Psychological Bulletin*, 117(3), 497-529.
- Bekoff, M. y Pierce, J. (2009). *Wild Justice: The Moral Lives of Animals*. University of Chicago Press.
- Brosnan, S. F. y De Waal, F. B. M. (2003). Monkeys Reject Unequal Pay. *Nature*, 425, 297-299.
- Camerer, C. F. y Thaler, R. H. (1995). Anomalies: Ultimatums, Dictators and Manners. *Journal of Economic Perspectives*, 9(2), 209-219.
- Coe, C. L. y Rosenblum, L. A. (1984). Male Dominance in the Bonnet Macaque: A Malleable Relationship. P. R. Barchas y S. P. Mendoza

- (eds.), *Social Cohesion: Essays Toward a Sociophysiological Perspective*. (pp. 31–64). Greenwood Press.
- Cummins, R. (1975). Functional Analysis. *The Journal of Philosophy*, 72(20), 741–765.
- Cummins, R. y Roth, M. (2010). Traits Have Not Evolved to Function the Way They Do Because of a Past Advantage. En F. J. Ayala y R. Arp (eds.), *Contemporary Debates in Philosophy of Biology*. (pp. 72–85). Blackwell.
- Davies, P. S. (2009). Conceptual Conservatism: The Case of Normative Functions. En U. Krohs y P. Kroes (eds.), *Functions in Biological and Artificial Worlds*. (pp. 127–146). MIT Press.
- Darwin, C. R. ([1882] 1989). *The Descent of Man, and Selection in Relation to Sex*. New York University Press.
- De Waal F. B. M. (1982). *Chimpanzee Politics: Power and Sex Among Apes*. Jonathan Cape.
- De Waal, F. B. M. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Harvard University Press.
- De Waal, F. B. M. (2007). Los instintos sociales de los primates, la moralidad humana y el auge y la caída de la “teoría de la capa”. En *Primates y filósofos. La evolución de la moral del simio al hombre*. (pp. 23–105). Paidós.
- Flack, J. C. y De Waal, F. B. M. (2002). ‘Any Animal Whatever’. Darwinian Building Blocks of Morality in Monkeys and Apes. *Journal of Consciousness Studies*, 7(1-2), 1-29.
- Fodor, J. A. (1983). *The Modularity of Mind*. MIT Press.
- Friedman, B. M. (2005). *The Moral Consequences of Economic Growth*. Vintage Books.
- Gausel, N. y Leach, C. (2011). Concern for Self-Image and Social Image in the Management of Moral Failure: Rethinking Shame. *European Journal of Social Psychology*, 41, 468-478. DOI: <https://doi.org/10.1002/ejsp.803>.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.
- Gould, S. J. y Vrba, E. S. (1982). Exaptation—A Missing Term in the Science of Form. *Paleobiology*, 8(1), 4-15
- Greenfield, M. D. (2006). Honesty and Deception in Animal Signals. En J. R. Lucas y L. W. Simmons (eds.), *Essays in Animal Behaviour: Celebrating 50 Years of Animal Behaviour*. (pp. 281-300). Elsevier Academic Press.

- Haidt, J. y Joseph, C. (2004). Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues. *Daedalus*, 133(4), 55–66.
- Hamilton, W. D. (1964). The Genetical Theory of Social Behaviour. II. *Journal of Theoretical Biology*, 7(1), 17-52.
- Harms, W. F. y Skyrms, B. (2008). Evolution of Moral Norms. En M. Ruse (ed.), *The Oxford Handbook on the Philosophy of Biology*. (pp. 434-450). Oxford University Press.
- Haun, D. B. M. y Tomasello, M. (2011). Conformity to Peer Pressure in Preschool Children. *Child Development*, 82(6), 1759–1767. DOI: <https://doi.org/10.1111/j.1467-8624.2011.01666.x>.
- Huneman, P. (2013). *Functions: Selection and Mechanisms*. Springer.
- Jackson, T. N. W. y Fry, B. G. (2016). A Tricky Trait: Applying the Fruits of the “Function Debate” in the Philosophy of Biology to the “Venom Debate” in the Science of Toxinology. *Toxins*, 8, 263.
- James, H. (2002). *The End of Globalization: Lessons from the Great Depression*. Harvard University Press.
- Jamieson, D. (2002). Is There Progress in Morality? *Utilitas*, 14(3), 318-338.
- Joyce, R. (2006). *The Evolution of Morality*. MIT Press.
- Kitcher, P. (2006). Biology and Ethics. En D. Copp (ed.), *The Oxford Handbook of Ethical Theory*. (pp. 163-186). Oxford University Press.
- Krebs, J. R. y Dawkins, R. (1984). Animal Signals: Mind-Reading and Manipulation. En J. R. Krebs y N. B. Davies (ed.), *Behavioural Ecology: An Evolutionary Approach*. (pp. 380-402). Blackwell.
- Lewis, D. (1969). *Convention. A Philosophical Study*. Harvard University Press.
- Machery, E. y Mallon, R. (2010). Evolution of Morality. En J. M. Doris y The Moral Psychology Research Group (eds.), *The Moral Psychology Handbook*. (pp. 3–47). Oxford University Press.
- Maslow, A. H. (1968). *Toward a Psychology of Being*. Van Nostrand.
- Maynard Smith, J. y Harper, D. (2003). *Animal Signals*. Oxford University Press.
- Millikan, R. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. MIT Press.
- Neander, K. (1991). Functions as Selected Effects: The Conceptual Analysts Defense. *Philosophy of Science*, 58(2), 168–184.

- Over, H. (2016). The Origins of Belonging Social Motivation in Infants and Young Children. *Philosophical Transactions of the Royal Society B*, 371. DOI : <https://doi.org/10.1098/rstb.2015.0072>.
- Peterson, D. (2011). *The Moral Lives of Animals*. Bloomsbury Press.
- Prinz, J. J. (2007). *The Emotional Construction of Morals*. Oxford University Press.
- (2008). Is Morality Innate? En W. Sinnott-Armstrong (ed.), *Moral Psychology. Volume 1. The Evolution of Morality: Adaptations and Innateness*. (pp. 367-406). MIT Press.
- Rizzolatti, G. y Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, 27, 169-192.
- Rogers, C. (1961). *On Becoming a Person*. Houghton Mifflin.
- Rosas, A. (2005). La moral y sus sombras: la racionalidad instrumental y la evolución de las normas de equidad. *Crítica. Revista Hispanoamericana de Filosofía*, 37(110), 79-104.
- Ruse, M. y Wilson, E. O. (1986). Moral Philosophy as Applied Science. *Philosophy*, 61(236), 173-192.
- Sartre, J.-P. (1946). *L'esistenzialismo è un umanismo*. Les Éditions Nagel.
- Schacter, S. (1959). *The Psychology of Affiliation. Experimental Studies of the Sources of Gregariousness*. Stanford University Press.
- Smetana, J. G. y Braeges, J. L. (1990). The Development of Toddlers' Moral and Conventional Judgments. *Merrill-Palmer Quarterly*, 36(3), 329-346.
- Sinclair, N. (2012). Metaethics, Teleosemantics and the Function of Moral Judgements. *Biology and Philosophy*, 27(5), 639-662.
- Singer, P. (2011). *The Expanding Circle: Ethics, Evolution, and Moral Progress*. Princeton University Press.
- Slocombe, K. E. y Zuberbühler, K. (2007). Chimpanzees Modify Recruitment Screams as a Function of Audience Composition. *Proceedings of the National Academy of Sciences of the United States of America*, 104(43), 17228-17233.
- Street, S. (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies*, 127, 109-166.
- Tajfel, H. (1981). Human Groups and Social Categories. Cambridge University Press [(1984). *Grupos humanos y categorías sociales*. Herder].
- Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 46(1), 35-57.

- Von Rohr, C. R., Burkart, J. M. y Van Schaik, C. P. (2011). Evolutionary Precursors of Social Norms in Chimpanzees: A New Approach. *Biology and Philosophy*, 26, 1–30.
- Wong, D. B. (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.
- Wright, L. (1973). Functions. *The Philosophical Review*, 82(2), 139–168.