

# LA OTRA EVOLUCIÓN:

## Darwin, la conciencia y el comportamiento moral

Pablo Quintanilla

(1) Quisiera comenzar agradeciendo a los organizadores de *Intersapiens* por esta amable invitación, y también especialmente a Susana y a Eric por su generosa disposición para intercambiar ideas sobre estos temas que nos interesan a todos. Me siento sumamente honrado por esta oportunidad.

El futuro del conocimiento es transdisciplinario, y encuentros como el de hoy nos permiten acostumbrarnos a salir del canon y de los límites de nuestras propias disciplinas para aprender de las vecinas. Esta tarea es riesgosa y difícil, pero cada vez más necesaria en el mundo académico actual. Es, además, una labor paradójica, porque aunque las ciencias tienden a la hiper especialización, cada vez resulta más claro que muchos aportes al conocimiento se dan en las intersecciones entre las disciplinas; precisamente porque los límites entre las ciencias son artificiales y, en el fondo, administrativos, no son contornos que reflejen divisiones de la realidad en sí misma.

Por eso, la ponencia que escucharán hoy no será una clásica conferencia de filosofía, sino, siguiendo el sentido mismo de *Intersapiens*, un ensayo transdisciplinario en que voy a intentar navegar, con bandera filosófica, entre mares igualmente turbulentos pero distintos en sus contenidos y particularidades. Así, entonces, *la primera idea* que deseo transmitir es que, en el mundo académico de hoy, hay temas que ya no se pueden abordar desde una sola disciplina. *La segunda idea* es que el abismo que solía separar a las ciencias naturales de las humanas se ha ido reduciendo, no porque uno de estos dos grupos de ciencias vaya absorbiendo al otro, sino porque se han

superpuesto, donde el espacio de la intersección se hace cada vez más elaborado y profundo, enriqueciéndose en complejidad.

**(2)** Eso es precisamente lo que ocurre con el tema que abordaremos esta noche, pues se encuentra en la intersección entre la filosofía de la mente, la ética y la psicología moral; y es el producto de la aplicación de algunas tesis centrales de la teoría de la evolución a la pregunta por el origen y el desarrollo de ciertas funciones psíquicas humanas y del comportamiento moral. **(3)** Lo que nos vamos a preguntar es cómo y por qué surgió el comportamiento moral en nuestra especie, y si esa evolución está asociada con la evolución de otros procesos cognitivos y afectivos.

**(4)** La idea central, que quisiera plantear para la discusión, es que el comportamiento moral humano no es solo un producto cultural, como solía creerse por mucho tiempo, sino también está enraizado en bases biológicas que tienen que ver con las necesidades de supervivencia de la especie. **(5)** En otras palabras, ciertas formas de comportamiento que están a la base de la conducta moral, como la cooperación solidaria, el altruismo y la compasión tendrían una función adaptativa en una especie en la que la supervivencia del individuo depende crucialmente de su pertenencia a un grupo con complejas relaciones sociales. **(6)** La idea es también que hay correlaciones entre estas formas de comportamiento y ciertas funciones cognitivas, especialmente la capacidad de simulación, que es la habilidad de atribuir estados mentales como creencias, deseos y afectos a otros individuos, para así poder comprender y predecir su comportamiento. Es posible sostener entonces que esas funciones cognitivas coevolucionaron con formas de comportamiento que están a la base del comportamiento moral.

**(7)** Presentaré lo que se suele llamar una visión “multicomponential” del comportamiento moral. Según esta, la moral sería el producto de muchos componentes diferentes (de tipo cognitivo, afectivo y social) cada uno de ellos con valor adaptativo, los que al superponerse en nuestra especie habrían generado un complejo fenómeno totalmente nuevo en la naturaleza: la moral, el

valor, la normatividad y la experiencia del deber. La moral sería, así, un *by product* o un producto colateral de esos componentes previos.

Es importante notar, sin embargo, que la presencia de estas formas de comportamiento y funciones cognitivas sería condición necesaria pero no suficiente para la presencia de comportamiento moral en un sentido pleno. **(8)** Un individuo se convierte en un agente moral, en sentido estricto, solo cuando está en condiciones de elaborar juicios morales con pretensiones de objetividad y cuando rige sus acciones a partir de normas que él o ella reconoce como valiosas y que ha elegido autónomamente. El punto, sin embargo, es que para que se pueda dar ese fenómeno moral en sentido pleno es necesario que se haya producido toda una historia evolutiva sin la cual no hubiera sido posible. Es esa historia evolutiva la que nos interesa ahora. Comenzaremos entonces con un tema que está en la frontera entre la biología y la filosofía moral: el altruismo. **(9)** ¿Cómo es posible el comportamiento altruista en un mundo natural donde *aparentemente* rige como ley suprema la lucha por la supervivencia? **(10)** (Esa es la lucha por la supervivencia)

### I ¿Cómo es posible el altruismo?

Las discusiones en biología teórica y filosofía moral han producido abundante bibliografía en relación a la naturaleza, génesis y posibilidad del altruismo. **(11)** Desde el punto de vista biológico, un individuo es altruista si tiene un comportamiento que mejora las posibilidades reproductivas o de supervivencia de otros individuos, a expensas de las suyas propias. La existencia de este fenómeno presenta, sin embargo, algunas preguntas. Todos los seres vivos estamos programados para luchar por nuestra supervivencia o, en todo caso, por la de los individuos que comparten nuestros genes. Sin embargo, con frecuencia nos comportamos de manera altruista tanto con quienes comparten nuestros genes como con quienes no los comparten. Esto no solo ocurre entre seres humanos sino también entre algunas especies de animales. ¿Por qué ocurriría esto? **(12)** Aquí aparece la denominada *paradoja del altruismo*: si el comportamiento altruista es un rasgo hereditario, y dado que

los individuos altruistas se ponen a sí mismos en circunstancias menos ventajosas que los egoístas, sería de esperar que la selección natural y la presión del entorno vayan eliminando progresivamente a los altruistas hasta su total desaparición.

**(13)** La teoría clásica para resolver esta paradoja procede de los capítulos cuarto y quinto de *El origen del hombre*<sup>1</sup> de Darwin, y suele recibir el nombre de selección de grupos (*group selection*). Según esta teoría son los grupos más cooperativos entre sí los que tienen mejores posibilidades de supervivencia respecto de grupos menos cooperativos, con lo cual la lucha por la supervivencia seleccionaría el rasgo altruista que a su vez se transmitiría hereditariamente. **(14)** (Esa es la carátula de la primera edición de *El origen del hombre*) **(15)** El problema que Darwin no resolvió es cómo llegan los individuos más cooperativos a sobrevivir frente a la competencia de individuos egoístas, antes que se conforme el grupo cooperativo. Para resolver ese problema surgieron nuevas propuestas. **(16)** William Hamilton propuso la teoría de la selección de parentesco (*kin selection*),<sup>2</sup> según la cual la naturaleza ha seleccionado individuos altruistas porque esto favorece la supervivencia de las familias, es decir de los genes compartidos. **(17)** Robert Trivers propuso la teoría del altruismo recíproco,<sup>3</sup> según la cual es ventajoso para un individuo realizar actos altruistas pues estos generarán en los demás miembros del grupo actos igualmente altruistas respecto de ese individuo. **(18)** Según Richard Dawkins, en su famoso libro *El gen egoísta*, quien aspira a la supervivencia no es el individuo sino sus genes, con lo cual estos nos “programan” para beneficiar a aquellos con quienes compartimos genes, siendo nosotros en realidad, por así decirlo, “vehículos de transmisión genética”. Posteriormente, en 1998, se produjo un renacimiento de la teoría de la selección de grupos gracias a Sober y Wilson,<sup>4</sup> pero son muchos los autores

---

<sup>1</sup> Cf. Darwin, Charles, *El origen del hombre y la selección en relación al sexo*, México, Editorial Diana, 1969.

<sup>2</sup> Hamilton, W. D. (1964), “The genetical evolution of social behaviour”, en: *Journal of Theoretical Biology*, 7.

<sup>3</sup> Cf. Trivers, R. (1971), “The evolution of reciprocal altruism”, en: *Quarterly Review of Biology*, 46.

<sup>4</sup> Cf. Sober, E. y Wilson, D.S., *Unto Others: The Evolution of Psychology of Unselfish Behaviour*, Cambridge, MA: Harvard University Press, 1998; Wilson, D.S., “Hunting, sharing

que han aportado elementos a una concepción más comprehensiva de la génesis del altruismo que incluye rasgos de más de una de estas teorías. Muchas de ellas, sin embargo, parecen presuponer algo discutible y es que las motivaciones esenciales del comportamiento humano son fundamentalmente egoístas y, por tanto, hay que explicar el altruismo como una excepción de ese egoísmo que nos gobernaría por *default*. **(19)** En esto, esas posiciones están en la línea de Thomas Hobbes para quien uno se comporta moralmente porque cree que al largo plazo eso terminará beneficiándolo.

Lo que es interesante es que recientemente diversos autores han objetado ese presupuesto fundamental, el carácter egoísta del hombre, y han propuesto la existencia de una tendencia altruista –tan poderosa como el egoísmo- que se explicaría por el carácter social de la especie y por el desarrollo de lo que se suele llamar una “inteligencia social” en los primates superiores. En esta línea están autores como **(20)** Frans de Waal, **(21)** Peter Singer, Kim Sterelny, Richard Joyce, **(22)** Marc Hauser, Dorothy Cheney y Robert Seyfarth, o **(23)** Walter Sinnott-Armstrong. (Es interesante que estos autores pertenecen a distintas especialidades: antropología, ciencias cognitivas, filosofía, psicología, biología y etología).

Desde estas distintas ciencias estos autores sostienen que no somos brutos egoístas barnizados por una leve capa de moralidad cultural que reprime nuestro egoísmo natural, sino más bien que nuestras tendencias sociales son tan poderosas como las egoístas y tienen el mismo origen biológico adaptativo. Es tan beneficioso para la supervivencia de la especie ser egoísta como ser altruista **(24)** Así, en vez de darle la razón a Hobbes parecería que los enfoques transdisciplinarios se estarían aproximando a una concepción semejante a la de Aristóteles.<sup>5</sup>

Al parecer, la clave para entender este proceso es estudiar la evolución de ciertas funciones cognitivas. **(25)** Así, nos preguntaremos ahora cómo

---

and multilevel selection: the tolerated theft model revisited”, *Current Anthropology*, 39, (1): 73-97, 1998.

<sup>5</sup> De este lado tienden a estar, por ejemplo, de Waal, Franz y Singer, Peter. Cf. De Waal, *Primates and Philosophers. How Morality Evolved*, Princeton University Press.

puede el estudio de la evolución de las funciones cognitivas, especialmente la capacidad de simulación, ayudar a entender la evolución del altruismo y de la moralidad. Sin duda no se está asumiendo que esta sea la *única* explicación del origen del comportamiento moral, sino una explicación más de este fenómeno. Esta es la diferencia entre un naturalismo reductivista y uno no reductivista. El primero sostiene que la *única* explicación válida de todos los fenómenos, incluyendo los humanos, es la que puede reducirse a relaciones causales naturales; la segunda considera que la explicación en términos de relaciones causales naturales, siendo importante, no es la única válida. Esta distinción es esencial para mostrar la compatibilidad entre distintos tipos de explicaciones. Darwin mismo era un naturalista no reductivista y lamentaba algunas interpretaciones reductivistas que se hicieron de su obra.

Vamos, entonces, a concentrarnos ahora en la capacidad de simulación como una de las funciones cognitivas que estarían en la base del comportamiento moral. **(26)** Como ya hemos visto, simulación es la capacidad de imaginar los estados mentales ajenos, en condiciones contrafácticas, como si fueran propios. **(27)** Solo hay un matiz de diferencia pero no vamos a entrar ahora en los detalles, con el concepto de metarrepresentación, entendido como la capacidad para representarse los estados mentales ajenos. Para los fines de esta conversación vamos a asumir que simulación y metarrepresentación se dar juntos.

El punto es que estas capacidades tienen una alta función de supervivencia porque permiten predecir el comportamiento de otros individuos, lo que era central para nuestros antepasados en comportamientos cooperativos como la defensa y la caza. Para poder interactuar exitosamente con otras personas es necesario tener la capacidad de atribuirles los estados mentales que nosotros creemos que tendríamos si estuviéramos en las circunstancias en que creemos están ellos. De esa manera no solo imaginamos lo que pasa por sus mentes actuales sino también lo que pasaría si nosotros actuáramos de una u otra forma. Al hacerlo, podemos predecir sus acciones y planificamos las acciones que nosotros realizaríamos en caso que ellos actuaran como nosotros estamos previendo. Es como un juego de ajedrez en que nos adelantamos a

posibles jugadas de nuestros rivales, imaginando distintos escenarios posibles que involucran las acciones ajenas y las reacciones nuestras, así como las nuevas reacciones de los demás. Estas habilidades hacen posible la transmisión de información y la construcción de un orden social que tiene complejas relaciones de jerarquía y poder, así como de solidaridad, reciprocidad y altruismo. La correlación entre complejidad social y capacidad metarrepresentacional sugiere que las habilidades cognitivas y afectivas humanas, que son las que participan en las capacidades metarrepresentacional y de simulación, han evolucionado modeladas según su función social.<sup>6</sup> Esto es lo que se suele llamar la hipótesis de la inteligencia social o del cerebro social. Según el antropólogo Robin Dunbar, las investigaciones neurológicas del neocórtex sugieren la hipótesis de que ya en los primates existiría un “cerebro social”, es decir, una zona del cerebro especializada en la predicción y coordinación del comportamiento social.<sup>7</sup>

Como se sabe, la ventaja adaptativa de los primates es su comportamiento gregario y organizado, pues individualmente son animales frágiles y vulnerables. **(28)** Sin embargo, dado que la convivencia interindividual genera inevitablemente problemas de coordinación y acuerdo, la selección natural habría favorecido la supervivencia de grupos e individuos con habilidades sociales de cooperación y manejo de conflictos, aunque también de manipulación y creación de alianzas para obtener el control del grupo. Los grupos de primates normalmente se mantienen unidos contra grupos rivales de la misma tropilla, pese a que contienen subgrupos que son rivales entre sí. Estos subgrupos son el producto de alianzas que se van modificando según el mayor o menor poder que adquieran los individuos que los conforman. (Las comparaciones con la vida política humana, por supuesto, no son pura coincidencia). **(29)** En todo caso, el punto es que la planificación y la coordinación social requieren de formas de razonamiento que nos permiten

---

<sup>6</sup> Cf. Cosmides, L. (1989), “The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task”, en: *Cognition*, 31: 187-276; Byrne, R.W., y Whiten A. (eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*, Oxford: Clarendon Press, 1988.

<sup>7</sup> Cf. Dunbar, R.I.M., (1998), “The social brain hypothesis”, en: *Evolutionary Anthropology*, 178-190; Obiols, J., y Pousa, E., “La teoría de la mente como módulo cerebral evolutivo”, en: Sanjuán, J., Cela-Conde, C.J., *La profecía de Darwin. Del origen de la mente a la psicopatología*, Barcelona: Ars Médica, 2005.

*simular* escenarios futuros en los que podríamos estar involucrados. **(30)** El imaginar esas situaciones posibles y elegir entre ellas nos permite probar por anticipado el éxito o fracaso de nuestras decisiones.

**(31)** Se ha propuesto la existencia de lo que se denomina un *simulador interno* que permitiría al agente imaginar escenarios alternativos realistas para poder actuar exitosamente. Esta capacidad también haría posible que el individuo delibere, realizando cálculos acerca de las mejores consecuencias de sus actos en términos del mayor beneficio individual y grupal. En términos morales, la atribución de estados mentales a otros individuos con los cuales uno se sentiría moralmente comprometido comenzaría con los individuos genética y socialmente más cercanos a uno y se ampliaría, como el círculo en expansión de Peter Singer<sup>8</sup> y la torre de la moralidad de Frans de Waal,<sup>9</sup> hasta incluir otras o incluso todas las criaturas vivas.

## II Altruismo y comportamiento moral

Desde un punto de vista evolutivo, podríamos describir cuatro niveles en la evolución del comportamiento moral, donde cada uno de ellos está presente en el nivel posterior.

**(32)** En primer lugar tenemos el *comportamiento cooperativo*: Este está caracterizado por ser un tipo de ayuda, colaboración o asistencia que se realiza para obtener un fin colectivo que beneficia al grupo. Hay comportamiento cooperativo en todos los animales sociales, incluidos los insectos. **(33)** En segundo lugar tenemos el *altruismo biológico*: Aquí se trata de un tipo de colaboración que beneficia las posibilidades de supervivencia o reproducción de otro individuo a expensas de las de uno mismo. Este tipo de altruismo se puede encontrar en la mayor parte de animales sociales. **(34)** En tercer lugar está el *altruismo moral*: En este caso se trata de una colaboración libre,

---

<sup>8</sup> Cf. Singer, Peter, *The Expanding Circle*: Oxford, Clarendon Press, 1981.

<sup>9</sup> Cf. De Waal, Frans, *Bien natural*, Barcelona: Herder, 1997. (*Good Natured: The Origins of Right and Wrong in Humans and Other Animals*, Cambridge MA: Harvard University Press, 1996); *Primates and Philosophers. How Morality Evolved*, Princeton University Press, 2006.

*consciente*, intencional y desinteresada que beneficia a otro individuo (con el que se puede o no compartir carga genética) a expensas de los intereses de uno mismo, donde el bienestar del otro es considerado un fin en sí mismo. **(35)** Finalmente tenemos el *comportamiento moral*. Este tiene los rasgos del altruismo moral, pero, además, en este caso el agente considera que sus acciones están regidas por normas morales que él ha elegido voluntariamente y con las que desea regir su comportamiento, es decir, tiene concepto del deber, lo que le permite elaborar juicios morales. Se verá, entonces, que puede haber altruismo moral sin que haya noción de deber; esto ocurre, por ejemplo, cuando el agente actúa voluntariamente por compasión. Pero incluso en el caso del altruismo moral ya debe estar presente cierta capacidad de deliberación. **(36)** Sin embargo, para que el agente actúe moralmente en sentido pleno, debe creer que se encuentra obligado para actuar de esa manera por normas que él o ella ha elegido o aceptado libremente, lo que le permite formular juicios morales.

**(37)** Ahora bien, intencionalidad, deliberación, conciencia de las motivaciones, libre albedrío y autonomía también son un asunto de grado y forman una continuidad. Es decir, en general podemos atribuir estas propiedades a las acciones realizadas por seres humanos adultos, pero en tanto vamos retrocediendo en la vida de la persona hasta su infancia (o en la evolución de nuestra especie desde otros homínidos), reconoceremos cada vez menos de estos rasgos en su comportamiento. No todo acto intencional es plenamente consciente (considérese los actos fallidos y los *lapses* freudianos) y no todo acto consciente es plenamente intencional (a veces actuamos manipulados o bajo presión).<sup>10</sup> Pero conciencia e intencionalidad van de la mano, incluso si en proporciones diferentes. En cuanto a la continuidad, la capacidad de deliberar normativamente acerca de las consecuencias de nuestras acciones también es un asunto de grado, con lo cual es razonable suponer que lo es el comportamiento moral en general, teniendo como criterio básico de comparación el comportamiento moral humano. En su versión

---

<sup>10</sup> Si bien todo acto es por definición intencional, la intención puede ser consciente o inconsciente, en un sentido psicoanalítico. Uno, por ejemplo, podría realizar una acción motivado por causas que no conoce, pero seguiría siendo una acción porque habría habido intencionalidad, aunque fuera inconsciente.

humana desarrollada, desde por lo menos el altruismo moral, estas capacidades parecen inseparables de la capacidad de simulación; en tanto esta capacidad está menos presente, también lo está el altruismo y el comportamiento moral. Vamos a precisar un poco más esta conexión.

### **III Simulación y metarrepresentación**

La teoría más aceptada acerca de la evolución cognitiva es el modelo de autocatálisis. Según este, cuando los homínidos de hace aproximadamente dos millones de años se volvieron bípedos, como consecuencia de cambios en el ambiente, sus extremidades anteriores quedaron liberadas, con lo cual pudieron ser usadas para la confección y uso de artefactos. El uso de utensilios potenció la inteligencia y aumentó el volumen de masa encefálica, de aproximadamente 450cc hace tres millones de años (comparable a los cerebros actuales de chimpancés y gorilas), hasta aproximadamente 1,400cc en el homo sapiens sapiens actual. También se desarrolló y perfeccionó el comportamiento cooperativo para la caza y la protección mutua, lo que desarrolló más aún la inteligencia.

Como hemos visto, las habilidades involucradas en el comportamiento cooperativo para la caza y la protección mutua tienen que haber evolucionado de la mano con mecanismos de reconocimiento de las necesidades, emociones, creencias y objetivos ajenos. Estos mecanismos requieren de la aparición de capacidades metarrepresentacionales y simpáticas. **(38)** “Simpatía” es uno de los llamados sentimientos morales por David Hume y Adam Smith para denotar la capacidad de sentir en uno mismo los sentimientos de otras personas. **(39)**

Hay una distinción técnica entre simpatía y empatía en la que no voy a entrar. El punto importante ahora es que formas rudimentarias de la simpatía pueden ser encontradas en mamíferos superiores y en bebés muy pequeños, con incluso solo horas de vida, en lo que se conoce como “contagio

emocional".<sup>11</sup> **(40)** Esto se produce cuando un individuo reconoce estados mentales en otro, especialmente aquellos que involucran sufrimiento pero también placer o dicha, reproduciéndolos en sí mismo. Esta reproducción suele ser no consciente y pre-cognitiva. Un ejemplo de esto es cuando tenemos un grupo de recién nacidos y uno de ellos se pone a llorar; lo más probable es que todos los demás se contagien y rompan también a llorar, incluso si no tienen ninguna incomodidad. El contagio emocional tiene un alto valor de supervivencia porque permite reaccionar apropiadamente ante situaciones compartidas como, por ejemplo, con temor y huida ante una amenaza común. **(41)** Probablemente el sustrato más básico del contagio emocional sea el *motor mimicry* o imitación mecánica del movimiento corporal ajeno, también de manera no consciente y pre-cognitiva. Esto es lo que ocurre, por ejemplo, con una manada de alpacas que mecánicamente dejan lo que están haciendo y se ponen a mirar en cierta dirección, una vez que una de las alpacas ha actuado de esa manera. **(42)** *Motor mimicry* y contagio emocional son probablemente las raíces evolutivas de lo que después serán la simpatía y la simulación, que originalmente tendrían como objetivo permitir que los individuos se apoyen mutuamente en la satisfacción de necesidades y se protejan de peligros comunes.

En lo que sigue vamos a imaginar una secuencia evolutiva que conduciría al altruismo y comportamiento moral, donde cada paso anterior es condición de posibilidad del posterior. Los conceptos entre paréntesis son explicaciones o se dan de manera asociada a los que vienen inmediatamente antes de ellos. Después de exponer el esquema intentaré justificarlo.

**(43)** Comportamiento cooperativo, altruismo biológico, motor mimicry, contagio emocional → simpatía (compasión) → reconocimiento de la propia identidad en un espejo, conciencia de sí mismo, identificación con otro **(44)** → metarrepresentación/simulación (teoría de la mente, capacidad de mentalización, capacidad de atribuir estados mentales en segundo o más

---

<sup>11</sup> Cf. Hartfield, E., Cacioppo, T.T., Raspón, R.L., (1993), "Emotional contagion", en: *Current Directions in Psychological Science*, 2:96-99. Según estos autores el contagio emocional comienza como una tendencia automática a imitar y sincronizar con otra persona posturas y gestos, convergiendo posteriormente en el plano emocional.

grados), **(45)** capacidad de postergación → altruismo moral → deliberación, reconocimiento de intencionalidad moral, interpretación y juicio moral, autonomía → comportamiento moral.

Es posible sugerir algunas correlaciones entre la evolución cognitiva y moral del niño, así como entre la evolución cognitiva y moral de la especie. La evolución cognitiva del niño toma más o menos la siguiente forma: **(46)** A partir del año los niños consuelan a los demás. El que esto sea un comportamiento tan precoz sugiere su carácter innato y, por tanto, adaptativo y no cultural.

**(47)** Aproximadamente a los catorce meses los niños pueden involucrarse en situaciones de atención compartida (*joint attention*) con otras personas acerca de objetos comunes, hacen contacto visual y pueden seguir la dirección de las miradas ajenas. Así, por ejemplo, si el bebé ve que la madre mira hacia algún objeto distante con interés, el niño sigue la mirada de la madre y reacciona ante las reacciones de ella. Estos son los rudimentos de lo que, mucho más adelante, será la capacidad de triangular, sobre lo que volveremos en un instante.

El contagio emocional progresivamente se va convirtiendo en capacidad simpática y hace posible la capacidad de identificación y la conciencia de sí mismo, lo que suele considerarse que ocurre cuando el niño puede reconocerse en un espejo, entre los 18 y los 24 meses. El retraso de la capacidad de reconocerse en un espejo está relacionado con el retraso en la capacidad de atribuir estados mentales a otros individuos, lo que suele ocurrir en niños con retardo mental, autismo y daño en el lóbulo frontal.<sup>12</sup>

**(48)** Los filósofos de la mente y los científicos cognitivos suelen aceptar la distinción que formuló Antonio Damasio<sup>13</sup> entre conciencia nuclear y conciencia extendida. La primera incluye la experiencia fenoménica, por ejemplo del dolor, la percepción y las sensaciones, así como la localización.

---

<sup>12</sup> Cf. Hauser, Marc, *Moral minds. How nature designed our universal sense of right and wrong*, Nueva York: Harper Collins, 2006.

<sup>13</sup> Cfr. Damasio, Antonio, *The feeling of what happens. Body and emotions in the making of consciousness*, Harvest Book, 1999.

Esta está presente en bebés muy pequeños y animales superiores. **(49)** La conciencia extendida (también llamada por algunos autores autoconciencia) incluye metacognición, es decir, procesos cognitivos acerca de nuestros procesos cognitivos; la identidad personal, el self, (es decir el *sí mismo*), la experiencia de la temporalidad, es decir, el sentido del yo como procediendo de nuestro pasado y con expectativas acerca del futuro. También la auto comprensión en términos de significados y valoraciones. Puede haber conciencia nuclear sin conciencia extendida, pero no al revés. La aparición de la conciencia extendida es más o menos simultánea con la aparición de la capacidad metarrepresentacional, que a su vez coincide con la capacidad de triangulación. Aparentemente ningún animal tiene conciencia extendida y, si la tiene, como en el caso de los animales que se reconocen en el espejo y tienen capacidad metarrepresentacional, esta sería de forma muy primitiva.

Alrededor de los dos años, precisamente con el lenguaje, aparece la capacidad conocida como *pretend play* (o juego simbólico), es decir, la función simbólica. Por ejemplo, el niño jugará a que el sillón de la sala es un caballo o a que su muñeca es un bebé, pero empezará a reconocer que esto es solamente un juego. Así, por ejemplo, se llevaría un buen susto si el sillón con el que juega se pusiera a relinchar.

**(50)** También aproximadamente a los dos años de edad el niño desarrolla la capacidad de representarse el mundo, con lo cual comienza a diferenciar la realidad de la fantasía (aunque a esa edad todavía suelen confundir la realidad con el sueño). **(51)** Alrededor de los tres años se produce otro salto cualitativo: el niño desarrolla la capacidad metarrepresentacional, es decir, la capacidad de representarse las representaciones ajenas, con lo cual puede atribuir a otros agentes estados mentales diferenciándolos de los suyos propios mediante creencias en segundo grado. Una creencia en primer grado es, por ejemplo, “yo creo que está lloviendo”. Una creencia en segundo grado es “yo creo que Susana cree que está lloviendo”. Una creencia en tercer grado es “yo creo que Susana cree que Eric cree que está lloviendo”. Y así en adelante. Alrededor de los dos años los niños solo tienen creencias en primer grado; alrededor de los tres años ya pueden tener creencias en segundo grado

y cerca de los cinco o seis años pueden tener creencias en tercero o más grados. **(52)** Esto se puede medir con las denominadas *pruebas de la falsa creencia*, que determinan la capacidad que tiene un niño (o un animal) para atribuir a otras criaturas creencias diferentes a las que él tiene, es decir, creencias que el niño cree son falsas, saliendo así de su perspectiva e *imaginando la perspectiva* del otro. De esta manera, el niño va aprendiendo a reconocer y entender el punto de vista ajeno sin perder el propio.

Los niños de menos de tres años creen que todas las personas creen lo que ellos creen, incluso que todos ven lo que ellos ven. Es decir, no tienen la capacidad de reconocer que hay otros puntos de vista *igualmente correctos* acerca de los mismos objetos.<sup>14</sup> (Como sabemos, algunas personas se quedan así para toda la vida). Todos los niños sanos desarrollan estas capacidades más o menos a la misma edad, excepto los casos severos de autismo y de daño en el lóbulo frontal. Los niños autistas tampoco entienden oraciones metafóricas ni contrafácticas y no tienen la capacidad de jugar, es decir, de imaginar tener una vida diferente de la que viven.

Alrededor de los tres años los niños comienzan a desarrollar la capacidad de triangulación en que constituyen los conceptos de *yo, otros y mundo compartido* como conceptos inseparables que se interdefinen mutuamente, donde cada uno de ellos se comprende solo a través de los otros dos. Eso les permite comenzar a desarrollar los conceptos de verdadero y falso, así como la distinción entre realidad objetiva y apariencia. **(53)** Algunos

---

<sup>14</sup> Las pruebas de la falsa creencia funcionan de la siguiente manera. En un teatro de marionetas se tiene a los personajes A y B. A juega con una canica y la deja en una canasta de mimbre, para luego salir del escenario. B saca la canica de la canasta y la deja en una caja de cartón, después también sale del escenario. A entra nuevamente al escenario sin saber que B ha cambiado de lugar la canica. El experimentador pregunta a dos grupos de niños, un grupo de alrededor de 3 años y medio y otro grupo de cerca de 5 años, dónde buscará A la canica. Los niños más jóvenes contestarán que A la buscará en la caja de cartón (donde realmente está), mientras que los niños mayores dirán que A la buscará en la canasta de mimbre (donde A la puso). El experimento muestra que el grupo más joven carece aún de capacidad metarrepresentacional al no poder atribuir a A una creencia que ellos consideran falsa, pues eso supondría alejarse de sus propias creencias y ver el mundo desde otra perspectiva. Los niños mayores de cinco años, en cambio, no tienen ningún inconveniente en atribuir a A la creencia falsa y, por tanto, en representarse las representaciones de A. Este experimento prueba que ellos pueden representarse el mundo desde el punto de vista de A, sin perder su propia perspectiva de las cosas.

de los filósofos que más han investigado estos temas son Donald Davidson y Marcia Cavell.

Como hemos visto, mucho antes de desarrollar capacidades metarrepresentacionales los niños pequeños ya tienen vínculos afectivos con otras personas, lo que sugiere que la metarrepresentación se desarrolla desde habilidades preconceptuales no-conscientes que se van haciendo conscientes y desarrollando paulatinamente, a partir de factores neurológicos y de condiciones del entorno como el apego.

Los estudios<sup>15</sup> muestran interesantes correlaciones entre la capacidad metarrepresentacional y el juicio moral. **(54)** Por ejemplo, entre los cinco y los siete años los niños juzgan los actos moralmente, no según sus consecuencias objetivas sino dependiendo de las intenciones que los niños creen que tuvieron los agentes que las realizaron. Así, los niños juzgan de maneras diferentes actos que tuvieron exactamente las mismas consecuencias, si es que los niños atribuyen intenciones distintas a sus agentes. A esa edad los niños no solo atribuyen intenciones a las personas sino, además, les atribuyen intenciones que juzgan como buenas o como malas. Es claro que para hacer ese tipo de juicio moral los niños necesitan haber imaginado los estados mentales del agente (su buena o mala intención), lo que parece probar que el juicio moral presupone cierta capacidad de simulación.

El punto es que la capacidad de simulación hace posible la existencia de complejas relaciones sociales, entre estas está el reconocimiento del sufrimiento ajeno y una tendencia natural a aliviarlo, especialmente si se trata de un miembro del grupo más cercano. Este sería el origen del comportamiento altruista moral. Las investigaciones sugieren que, en condiciones particulares, el reconocimiento del sufrimiento ajeno produce un sufrimiento propio, lo que activa un comportamiento de cooperación. Existe este tipo de comportamiento en los seres humanos, aunque hay reportes documentados de comportamiento

---

<sup>15</sup> Cf. Hauser, M., *Moral minds*, op. cit., p.207.

compasivo en primates no humanos y en mamíferos como ballenas, delfines, elefantes, lobos y perros.<sup>16</sup>

**(55)** Las evidencias experimentales sugieren que hay una correlación entre capacidad metarrepresentacional, capacidad de postergación y altruismo.<sup>17</sup> Esto es fundamental porque el comportamiento moral normalmente requiere que uno postergue, o eventualmente anule, sus deseos inmediatos en favor de un beneficio o recompensa posterior, que puede ser un beneficio propio, el equilibrio social o un beneficio para otros. Además, tanto la capacidad de postergación como el altruismo requieren de la habilidad para imaginar estados mentales futuros. En el caso de la capacidad de postergación, imaginamos los estados mentales que tendríamos si postergáramos cierta situación presente a cambio de un beneficio futuro; en el caso del altruismo, imaginamos los estados mentales futuros de otra persona si actuáramos de cierta manera para con él. **(56)** Los niños que responden bien a las pruebas de la falsa creencia tienen una mayor tendencia a compartir con otros niños a cambio de una gratificación futura.<sup>18</sup>

Esto parece requerir que el individuo se imagine a sí mismo en un momento futuro, obteniendo o realizando para otros un beneficio que no se obtendría de no aceptar un perjuicio presente, o imaginando a otros gozando de un beneficio futuro que solo se daría si el agente hiciera un sacrificio presente. Mediante esta simulación el individuo no solo se imagina las posibles circunstancias del entorno, sino también los estados mentales propios y ajenos que se producirían de actuar de una manera o de otra. Ahora bien, lo esencial en el comportamiento moral no es solo ser capaz de comportamiento altruista, sino también hacerlo de manera voluntaria y como producto de algún tipo de deliberación moral, lo que incluye el reconocer intenciones (en uno mismo y en los demás) y valorarlas como buenas o malas. Esta capacidad parece

---

<sup>16</sup> Cf. De Waal, *Bien natural*, op. cit., cap. 2.

<sup>17</sup> Cf. Moore, C. y Macgillivray, S., (2004), "Altruism, prudence and theory of mind in preschoolers", en: *New Directions for Child and Adolescent Development*, 103, 51-62; Thompson, C., Barresi, J., Moore, C., (1997), "The development of future-oriented prudence and altruism in preschoolers", en: *Cognitive Development*, 12, 199-212.

<sup>18</sup> Cf. Hauser, op. cit., p. 217.

presuponer todas las capacidades mencionadas que los niños van adquiriendo progresivamente.

#### **IV Altruismo, simulación y metarrepresentación en primates no humanos**

Aunque es un hecho que los primates no humanos son capaces de comportamiento altruista biológico, no es posible atribuirles comportamiento moral en el sentido que venimos discutiendo, y es discutible si podemos atribuirles altruismo moral, pues eso implicaría atribuirles algún grado de capacidad metarrepresentacional y de simulación. Sin embargo, sí podemos encontrar en algunas especies las bases evolutivas de lo que en los humanos se convirtió en el altruismo moral.

Premack y Woodruff realizaron experimentos de la falsa creencia con chimpancés y publicaron los resultados en 1978, en un artículo clásico titulado “Does the chimpanzee have a theory of mind?”, en el que además acuñaron la expresión “teoría de la mente”.<sup>19</sup> **(57)** Aunque hay mucho debate al respecto, desde entonces los diversos experimentos sugieren que los chimpancés llegan a atribuir a otros individuos, ya sea humanos u otros primates, estados mentales diferentes de los de ellos mismos, hasta incluso un tercer grado de capacidad metarrepresentacional. Chimpancés, gorilas y bonobos tienen la habilidad de predecir, engañar, manipular y monitorizar el comportamiento de otros individuos, así como de crear alianzas en contra de otros. Esto fue denominado por Humphrey “inteligencia maquiavélica”.<sup>20</sup> Los chimpancés, además, tienen la habilidad de contraengañar a otros. El engaño es hacer creer a otro individuo algo que uno no cree. El contraengaño es hacerle creer al otro que uno cree algo que no cree. Esta habilidad requiere, por lo menos, de capacidad metarrepresentacional en tercer grado. Sin embargo estos animales se equivocan mucho cuando tienen que atribuir a distintos individuos diferentes

---

<sup>19</sup> Cf. Premack, D., y Woodruff, G. (1978), “Does the chimpanzee have a theory of mind?”, en: *Behavioural Brain Sciences*, 4: 515-526.

<sup>20</sup> Cf. Humphrey, N.K., “The social function of intellect”, en: Bateson, P.P.G. y Hinde, R.A. (eds.), *Growing points in ethology*, Cambridge: Cambridge University Press, 1976.

estados mentales, lo que muestra que si bien tienen capacidad de simulación esta es muy básica. Es de suponer que la capacidad metarrepresentacional humana plenamente desarrollada requiere de lenguaje, lo que sugiere que esta capacidad no tendría más de 50,000 años.

Por otra parte, chimpancés, gorilas y delfines pueden reconocer su propia imagen en un espejo, lo que podría ser un indicador de conciencia de sí mismo, es decir conciencia extendida pero en un sentido muy básico.<sup>21</sup> Sin embargo, en la medida que nos vamos alejando de los patrones evolutivos humanos, estas capacidades se hacen más tenues y diferenciadas. Por ejemplo, aunque los gorilas se reconocen en espejos y pueden monitorizar la mirada de otros individuos, no hay evidencia de que tengan éxito en la prueba de la falsa creencia. Algunos monos son capaces de distraer a sus rivales comportándose como si, y por tanto haciéndoles creer que, hay un depredador acercándose, pero no se reconocen en espejos ni superan pruebas de la falsa creencia. Byrne y Whiten<sup>22</sup> denominan a este comportamiento en monos “engaño táctico”, y consideran que es un comportamiento rutinario y mecánico que no presupone capacidad metarrepresentacional. De ser así, ello indicaría que no todo comportamiento social complejo requiere de capacidad metarrepresentacional y que las formas más primitivas de esa capacidad se encuentran en comportamientos mecánicos más básicos.

**(58)** Hay significativas correlaciones entre capacidad metarrepresentacional y comportamiento compasivo. La evidencia sugiere que la compasión está presente donde hay capacidad de simulación. Además, por lo menos en el caso de los niños, hay una correlación entre la capacidad de reconocerse a uno mismo en un espejo y la tendencia a compadecerse y consolar a quienes sufren. Esto sugeriría que el consuelo y la compasión

---

<sup>21</sup> Cf. Povinelli, D.J., y Preuss, T.M. (1995), “Theory of mind: evolutionary history of a cognitive specialization”, en: *Trends in Neurosciences*, 18: 418-424. Esto se prueba en condiciones experimentales poniéndole al animal un punto de color en la frente y mostrándole un espejo. El animal se tocará la frente reconociendo que la imagen reflejada es la de él mismo. Algunos animales también hacen gestos y piruetas e incluso se adornan la cabeza, coquetamente, con ramas y hojas mientras se miran al espejo.

<sup>22</sup> Cf. Byrne y Whiten (1988), “Tactical deception in primates”, en: *Behavioural and Brain Sciences*, 11: 233-273; “The emergence of metarepresentation in human ontogeny and primate phylogeny”, en: Whiten (ed.), *Natural theories of mind*, Oxford: Basil Blackwell, 1991.

requieren de conciencia de sí mismo y capacidad de imaginarse padeciendo lo que otro padece.<sup>23</sup>

Es posible que a la base de la capacidad de simulación se encuentre la simpatía, es decir, los lazos afectivos entre individuos de un mismo grupo. La simpatía, a su vez, reposaría sobre mecanismos más básicos aun, como el contagio emocional y el motor mimicry. De hecho, cuando uno observa una fotografía donde se ve a una persona sufriendo, por ejemplo con un rictus de dolor, la cara del observador tiende a imitar la de la imagen, y el observador reproduce en sí mismo la experiencia que supone tuvo el fotografiado. Así se genera un sentimiento inmediato de malestar, incomodidad, desagrado o, incluso, sufrimiento físico. Estas reacciones musculares son mecánicas y automáticas e, incluso, pueden ser no conscientes e invisibles al ojo no experimentado.<sup>24</sup> **(59)** El punto es que reconocer un estado mental en otro individuo implica, en algún sentido, compartirlo. No es necesario padecerlo plenamente pero sí imaginar cómo sería tenerlo, y esto es, de alguna manera, experimentar algo de él.

Todo esto sugeriría una secuencia evolutiva donde los primates menos evolucionados poseen contagio emocional y capacidad simpática. Los primates más cercanos a los humanos, los chimpancés, llegar a tener formas rudimentarias de capacidad de simulación. En el caso de los homínidos, la capacidad de simulación se habría potenciado con el desarrollo del cerebro y con la complejidad de las relaciones sociales, dando lugar a formas de altruismo más complejas. Finalmente, con la aparición del lenguaje, la capacidad metarrepresentacional humana se habría potenciado (y habría potenciado al lenguaje) al punto que habría conformado las bases para la aparición del comportamiento moral en su estado actual.

---

<sup>23</sup> Cf. Johnson, D.B. (1982), "Altruistic behaviour and the development of self in infants", en: *Merrill-Palmer Quarterly*, 28: 379-388; Bischof-Köler, D. (1988), "Über den Zusammenhang von Empathie und der Fähigkeit sich im Spiegel zu erkennen", en: *Schweizerische Zeitschrift für Psychologie*, 47: 147-159; De Waal, *Bien natural*, op. cit., p. 294, nota #34.

<sup>24</sup> Dimberg, U., Thunberg, M. y Elhemed, k. (2000), "Unconscious facial reactions to emotional facial expressions", en: *Psychological Science*, 11: 86-89.

**(60)** Para ir terminando y como una especie de resumen, se podría decir que el juicio moral, es decir la aprobación y desaprobación moral, solo está plenamente presente cuando hacemos juicios que tienen la pretensión de ser imparciales y generales.<sup>25</sup> **(61)** Esto ocurre cuando intentamos salir de nuestra propia perspectiva para ver el comportamiento y las intenciones, nuestras o ajenas, desde un punto de vista objetivo que tome en consideración las distintas perspectivas e intereses individuales. **(62)** Esto deber ir acompañado por el reconocimiento de que nuestro comportamiento está regido por normas que uno considera valiosas y que acepta voluntariamente. Solo en esas circunstancias podríamos hablar, de manera auténtica, de comportamiento moral, y es claro que este tipo de comportamiento solo se da en seres humanos.

**(63)** Hasta aquí hemos intentado explicar adaptativamente las bases evolutivas del comportamiento moral, pero todavía no resulta fácil explicar de la misma forma el salto del altruismo moral al comportamiento moral, es decir, a la aparición de normatividad y conciencia del deber. Hay gente que ha intentado hacerlo, pero no resulta inmediatamente claro que esos proyectos sean exitosos, ni si algún día llegarán a serlo.

**(64)** No se si alguna vez llegaremos a explicar adaptativamente toda la complejidad del comportamiento moral; lo que si creo que si lo hacemos será a partir del estudio de funciones cognitivas anteriores con valor adaptativo. **(65)** Pero esto será posible solo mediante un trabajo transdisciplinario. Mi impresión es que una disciplina sola difícilmente podrá hacerlo.

**(66)** ¡Muchas gracias!

---

<sup>25</sup> Cf. De Waal, *Primates and Philosophers*, op. cit., p. 144.