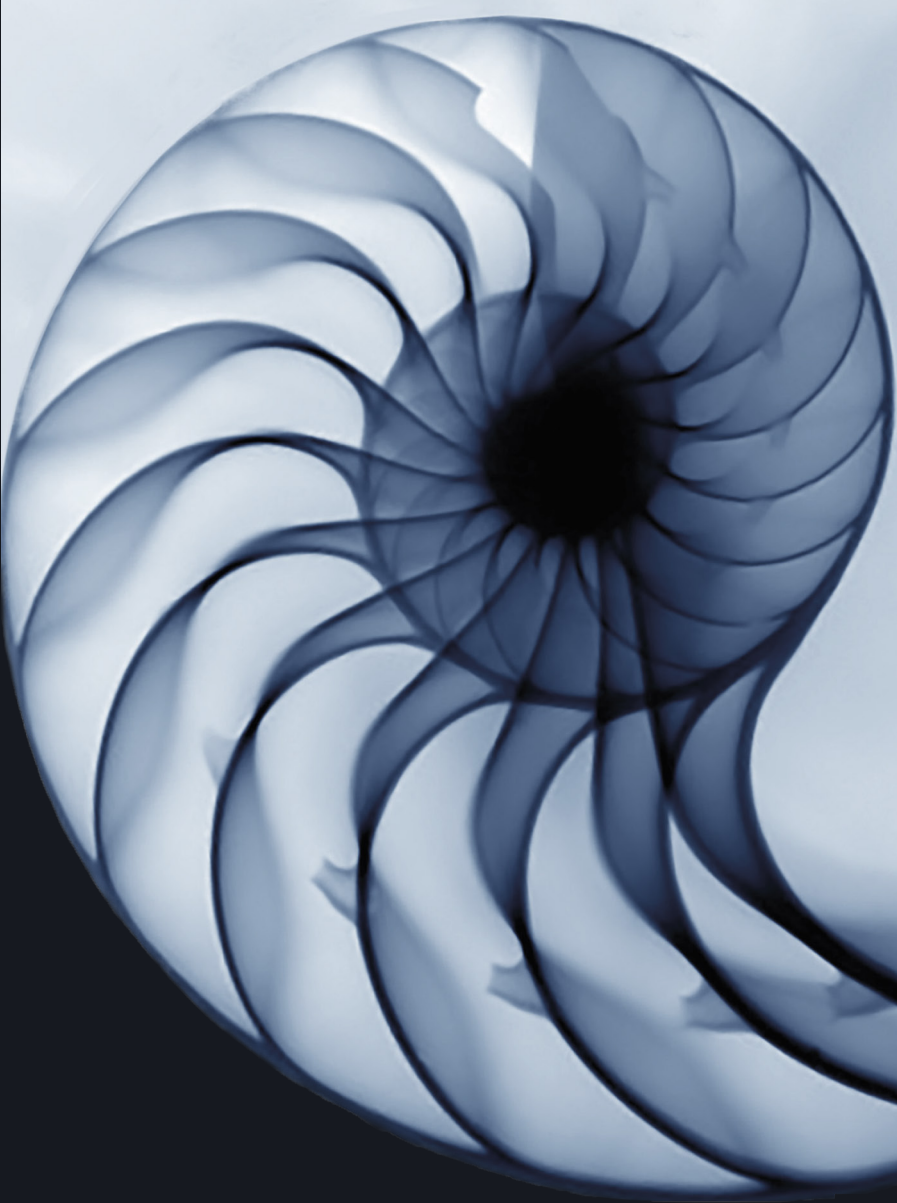


eISSN 1989-3612

# artefaCToS

Revista de estudios sobre la ciencia y la tecnología

Vol. 13, No. 1 (2024)



Ediciones Universidad  
Salamanca

INSTITUTO  
eCyt

CDU: 10 – IBIC: Filosofía de la ciencia (PDA) – BIC: Humanities / Philosophy / Philosophy: epistemology and theory of knowledge (HPK) – BISAC: Philosophy / Philosophy of Science / Philosophy & Social Aspects (PHI030000)

### Directores

M. Mar Cebrián Villar, Universidad de Salamanca, España  
Carina Gabriela Cortassa Amadio, Universidad Nacional de Entre Ríos / REDES - Centro de Estudios sobre Ciencia, Desarrollo y Educación Superior, Argentina  
Francisco Del Canto Viterale, Universidad de Dakota del Norte, Estados Unidos

### Editora

María Mar Cebrián Villar, Universidad de Salamanca, España

### Editores asociados

Esther Palacios Mateos, Universidad de Salamanca, España  
Benedicto Acosta Díaz, Universidad de Salamanca, España  
Mariano Martín Villuendas, Universidad de Salamanca, España

### Comité científico

Mario Albornoz, Universidad Nacional de General Sarmiento, Argentina  
Wiebe E. Bijker, Norwegian University of Science and Technology, Noruega  
Javier Echeverría, Ikerbasque, Basque Foundation for Science, España  
María Carmen Fernández Juncal, Universidad de Salamanca, España  
Santiago López García, Universidad de Salamanca, España  
Diana Maffía, Universidad de Buenos Aires, Argentina  
Carl Mitcham, Colorado School of Mines, Estados Unidos  
Carolina Moreno, Universidad de Valencia, España  
Miguel Ángel Quintanilla Fisac, Universidad de Salamanca, España  
Ana Rosa Pérez Ransanz, Universidad Nacional Autónoma de México, México  
José Manuel Sánchez Ron, Universidad Autónoma de Madrid, España  
Steven Shapin, Harvard University, Estados Unidos  
Peter Vermaas, Delft University of Technology, Países Bajos  
Hebe Vessuri, Universidad Nacional Autónoma de México, México

### Instituto de Estudios de la Ciencia y la Tecnología

Universidad de Salamanca  
Edificio I+D+i, Calle Espejo, n.º 2, 37007, Salamanca, España  
Teléfono: +34 923 294 834 – Email: artefactos@usal.es

### Ediciones Universidad de Salamanca

OJS - EUSAL Revistas  
Plaza de San Benito, n.º 2, 37008, Salamanca, España  
Teléfono: +34 923 294 598

ArtefaCToS se encuentra indexada en DOAJ, ERIH-PLUS, REDIB, DIALNET, Latindex 2.0, MIAR, EBSCO (Applied Science & Technology Source), Proquest (Aerospace Database, Agricultural & Environmental Science Database, Civil Engineering Abstracts, Metadex), cuenta con el sello de calidad FECYT.



**ArtefaCToS. Revista de estudios de la ciencia y la tecnología**

<https://revistas.usal.es/index.php/artefactos/index>

eISSN: 1989-3612 – DOI: <https://doi.org/10.14201/art2024131>

CDU: 10 – IBIC: Filosofía de la ciencia (PDA) – BIC: Humanities / Philosophy /  
Philosophy: epistemology and theory of knowledge (HPK) – BISAC: Philosophy / Philosophy of Science/  
Philosophy & Social Aspects (PHI030000)

**Vol. 13, No. 1 (2024), 2.<sup>a</sup> Época**

**Índice**

Introducción Frans B. M. DE WAAL	5-12
Introducción sección monográfica: “Primatología filosófica: reflexiones en torno a la obra de Frans de Waal” Joaquín SUÁREZ-RUIZ; Rodrigo LÓPEZ-ORELLANA	13-16
1. Normativity in the Wild. Insights from Frans de Waal Laurence KAUFMANN	17-25
2. ¿Normatividad primitiva en primates no humanos? Laura DANÓN	27-56
3. Justice and Related Matters in the Legacy of Frans de Waal Vicente CLARAMONTE SANZ; Rodolfo GUARINOS RICO	57-83
4. The Anthropocentric Bias in Animal Cognition Carolina SCOTTO	85-116
5. Evolved and Cultural Intuitions. Highly Speculative Remarks on the Origins of our Sense of Fairness Rodrigo BRAICOVICH	117-138
6. Frans de Waal, filósofo do direito Fábio PORTELA LOPES DE ALMEIDA	139-166
7. Mecanismos sociales, antropomorfismo y procesos cognitivos en animales no humanos. A 40 años de <i>La política de los chimpancés</i> Oscar David CAICEDO MACHACON; Rene J. CAMPIS; Eduardo BERMÚDEZ BARRER	167-208

8. Homología y antropomorfización crítica Julio TORRES MELÉNDEZ; Claudia MUÑOZ TOBAR	209-226
9. Del timing a la planificación: control temporal de la conducta en animales no-humanos Mauro ZAPATA	227-253
10. Theory-Construction in Comparative Cognition: Assessing the Case of Animal Normativity Nicolás SÁNCHEZ	255-277
11. Invasión silenciosa: la primatología de Imanishi y el sesgo cultural en la ciencia Frans B. M. DE WAAL	279-296
12. Normatividad natural: el “es” y el “debe” del comportamiento animal Frans B. M. DE WAAL	297-320

## INTRODUCTION

Frans B. M. DE WAAL

*C. H. Candler Professor Emeritus, Emory University, Atlanta, USA*

My relation with philosophy has always been somewhat ambiguous. Not because I believe science doesn't need philosophy, as some prominent scientists have recently argued. The love-hate relationship between empirical science and philosophy has never affected me. I have freely borrowed philosophical concepts to connect my primate observations with human behavior.

Instead, the ambiguity stems from the fact that I am no philosopher. My knowledge of philosophical texts, even those I have cited, remains limited. It doesn't reach the depth that real philosophers derive from extensive reading and a background in sorting out conflicting views. I am fully aware of this shortcoming, but can't resist philosophy anyway.

The habit to quote philosophers started with my work on the power struggles among chimpanzees, which I interpreted along Machiavellian lines. Even if many philosophers don't like to be associated with his ideas, Niccolò Machiavelli was a political philosopher. His impact was felt when my first book, *Chimpanzee Politics* (1982), which introduced the Florentine to primatology, led to (over)use of the term "Machiavellian Intelligence" to describe animal social cognition in general.

The philosophers who influenced me most, however, had little to say about political strategizing. My interests shifted around the 1980s when theoretical biologists began to depict the natural world as a dog-eat-dog place of combat, selfish genes, and devoid of kind intentions, as in Michael Ghiselin's oft-quoted line: "Scratch an 'altruist,' and watch a 'hypocrite' bleed." Genuine altruism couldn't exist. Moreover, animal cognition and emotion were never part of the discourse, nor was the love and attachment common in many social animals. Some of these writers sounded

utterly Cartesian, depicting animals as automatons (i.e. vehicles carrying out DNA programs). They presented these ideas to their readers as Darwinian even though Charles Darwin himself held very different views, and never classified animals as automatons. After all, Darwin wrote a whole book about animal and human emotions.

I found this whole cynical literature so depressing, and so contrary to the reality of observable animal and human behavior, that I read with great approval the works of philosophers such as Mary Midgley's, *Beast and Man* (1979), and Robert Richards' *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior* (1987). Both books presented a radically different view of nature, and sharply criticized the above writings by theoretical biologists. They also got me thinking about the origins of morality.

The connection with human morality came in handy when I got more and more involved in the study of animal empathy, reciprocity, and cooperation. My research automatically led me to explore what it was about human morality that made it radically different from animal behavior, as was often claimed. In the end, my conclusion was that the differences were overblown, and that it was better to adopt the Humean view that we are born with moral sentiments. Believing that these sentiments relate back to primate behavioral tendencies, I labeled them "building blocks" of morality. I never claimed, however, that a chimpanzee is a moral being in the same way that humans are. Instead, I advocated a connection between the way other primates behave and what in humans we call moral tendencies. These ideas were first explained in *Good Natured* (1996), and repeated and elaborated in subsequent books.

When we discovered a sense of fairness in monkeys, another moral building block could be added to the list (Brosnan & de Waal, 2003). Our discovery put a dent in *A Theory of Justice*, a famous treatise on this topic by American moral philosopher John Rawls (1972). In the grand Kantian tradition, Rawls tried to circumnavigate the emotions: "...for reasons of both simplicity and moral theory, I have assumed an absence of envy." Since when can we simply drop emotions from a discourse about human behavior? I find this baffling. The irony is that if there were zero envy in the world, no one would even care about fairness, because we'd never see a meaningful reaction to its absence. Rawls may call envy a "vice," but his whole treatise makes no sense without it. The enormous emotional investment humans make in rectifying unfairness and injustice -- the screaming protests, the marches, the violence, the endurance of police beatings -- are a forceful reminder that we aren't dealing with some bloodless mental

construct. Injustice shakes us to the core, something that no amount of abstract reasoning, such as that by Rawls, will ever accomplish.

And so I kept engaging with philosophical writing, either borrowing concepts or countering ideas I felt were divorced from reality. The latest such excursion was my writing about “normativity” (de Waal, 2014). There is a tendency to claim that human behavior is guided by social norms and expectations, whereas animal behavior just is. It’s the old Humean distinction between “is” and “ought.” There exists no agreement on this topic, which is why it remains a perennial of philosophical debate. Some have gone so far, however, as to wield “Hume’s guillotine” to kill off any and all attempts to apply evolutionary logic or neuroscience to human morality. Science can’t help us understand human morality, it is argued (Black, 1970).

In contrast, I believe that if capuchin monkeys reject a reward that they normally eat with gusto, only because their neighbor gets a much better reward for the same task, we are dealing with normativity. The monkey has an idea how rewards ought to be distributed. Chimpanzees go even further in that they may refuse the more preferred reward until their neighbor gets one, too. We found that chimpanzees play the Ultimatum Game in the same way as children, including a negative reaction to unfair outcomes. A child may shout “That’s unfair!” whereas a chimpanzee may spit water in the face of the one making an unacceptable proposal (Proctor *et al.*, 2013).

There are other examples of social normativity in animal behavior. If we take a broader view, normativity is everywhere. If I disturb a spider web, for example, the spider immediately works on repairing her web, which means that she has a model of what a web ought to look like. This counts as normativity, too, even if it isn’t concerned with social behavior. I feel that the topic of natural normativity deserves a much more extensive exploration than the one I provided.

The present volume takes a big leap in connecting philosophy with what we have learned about the behavior of our closest relatives. Instead of a biologist inspired by philosophy, the direction of contact is now the other way around. This is bound to lead to more dialogue. I notice, for example, how not every author agrees with my views and tries to salvage the age-old human-animal distinction that I have tried to undermine. A volume like this will also stimulate philosophers to pay more attention to the recent surge in animal cognition studies. This surge isn’t limited to the primates, but includes all sorts of species, also invertebrates. The sentience of all animals is now up for debate, and interaction between biologists and

philosophers will be crucial to resolve this issue (e.g. de Waal & Andrews, 2022).

I commend the editors for putting a volume like this together, and feel honored how it gives my life's work a central place.

## REFERENCES

- Black, M. (1970). *Margins of Precision: Essays in Logic and Language*. Ithaca, New York: Cornell University Press.
- Brosnan, S. F. & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, 425, 297-299. <https://doi.org/10.1038/nature01963>
- de Waal, F. B. M. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674033177>
- de Waal, F. B. M. (2007 [1982]). *Chimpanzee Politics: Power and Sex among Apes*. Baltimore, MD: Johns Hopkins University Press.
- de Waal, F. B. M. (2014). Natural normativity: The "is" and "ought" of animal behavior. *Behaviour*, 15, 185-204. <https://doi.org/10.1163/1568539X-00003146>
- de Waal, F. B. M. & Andrews, K. (2022). The question of animal emotions. *Science*, 375, 1351-1352. <https://doi.org/10.1126/science.abo2378>
- Midgley, M. (1979). *Beast and Man: The Roots of Human Nature*. London: Routledge. <https://doi.org/10.1038/281336a0>
- Proctor, D., Williamson, R. A., de Waal, F. B. M., & Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proceedings of the National Academy of Sciences USA*, 110, 2070-2075. <https://doi.org/10.1073/pnas.1220806110>
- Rawls, J. (1972). *A Theory of Justice*. Oxford: Oxford University Press.
- Richards, R. (1987). *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago: University of Chicago Press. <https://doi.org/10.7208/chicago/9780226149516.001.0001>



## INTRODUCCIÓN

Frans B. M. DE WAAL

*C. H. Candler Professor Emeritus, Emory University, Atlanta, USA*

Mi relación con la filosofía ha sido siempre un tanto ambigua. No porque crea que la ciencia no necesita de la filosofía, como han argumentado recientemente algunos científicos destacados. La relación de amor-odio entre la ciencia empírica y la filosofía nunca me afectó. He tomado prestados libremente conceptos filosóficos para conectar mis observaciones de primates con el comportamiento humano.

Más precisamente, la ambigüedad surge del hecho de que no soy un filósofo. Mi conocimiento de los textos filosóficos, incluso los que he citado, sigue siendo limitado. No alcanza la profundidad que los verdaderos filósofos obtienen de una lectura exhaustiva y de una experiencia en la resolución de puntos de vista conflictivos. Soy plenamente consciente de esta deficiencia, pero de todos modos no puedo resistirme a la filosofía.

La costumbre de citar filósofos comenzó con mi trabajo sobre las luchas de poder entre chimpancés, que interpreté en clave maquiavélica. Nicolás Maquiavelo fue un filósofo político, aunque a muchos filósofos no les gusta que se les asocie con sus ideas. Su impacto se dejó sentir cuando mi primer libro, *Chimpanzee Politics* (1982), que introdujo al florentino en la primatología, llevó a un uso (excesivo) del término 'inteligencia maquiavélica' en la descripción de la cognición social animal en general.

Sin embargo, los filósofos que más me influyeron tuvieron muy poco que decir sobre estrategia política. Mis intereses cambiaron alrededor de la década de 1980, cuando los biólogos teóricos comenzaron a describir el mundo natural como un lugar despiadado, con genes egoístas y desprovisto de buenas intenciones, como en la frase frecuentemente citada de Michael Ghiselin: "rasca un 'altruista' y verás sangrar a un 'hipócrita'". El altruismo genuino no podía existir. Además, la cognición y la emoción

animales nunca formaron parte de la discusión, ni el amor ni el apego eran vistos como comunes en muchos animales sociales. Algunos de estos escritores sonaban completamente cartesianos, describiendo a los animales como autómatas (es decir, vehículos que llevan a cabo programas de ADN). Presentaron estas ideas a sus lectores como darwinianas a pesar de que el propio Charles Darwin tenía puntos de vista muy diferentes y nunca clasificó a los animales como autómatas. Después de todo, Darwin escribió un libro completo sobre las emociones animales y humanas.

Toda esta literatura cínica me pareció tan deprimente y contraria a la realidad del comportamiento animal y humano observable, que leí con gran aprobación obras de filósofos como por ejemplo *Beast and Man* (1979), de Mary Midgley, y *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior* (1987), de Robert Richards. Ambos libros presentaron una visión radicalmente diferente de la naturaleza y criticaron duramente los escritos mencionados anteriormente de los biólogos teóricos. También me hicieron pensar en los orígenes de la moral.

La conexión con la moral humana me resultó útil cuando me involucré cada vez más en el estudio de la empatía, la reciprocidad y la cooperación animal. Mi investigación me llevó automáticamente a explorar qué era lo que hacía que la moral humana fuera radicalmente diferente del comportamiento animal, como a menudo se afirmaba. Al final, mi conclusión fue que las diferencias eran exageradas y que era mejor adoptar la visión humeana de que nacemos con sentimientos morales. Creyendo que estos sentimientos se relacionan con tendencias comportamentales de los primates, los denominé 'elementos básicos' de la moral. Sin embargo, nunca he afirmado que un chimpancé sea un ser moral del mismo modo que lo son los humanos. En cambio, abagué por una conexión entre la forma en cómo se comportan otros primates y lo que en los humanos denominamos 'tendencias morales'. Estas ideas fueron explicadas por primera vez en *Good Natured* (1996) y repetidas y elaboradas en libros posteriores.

Cuando descubrimos un sentido de justicia en los monos, se pudo agregar otro elemento moral a la lista (Brosnan y de Waal, 2003). Nuestro descubrimiento marcó una diferencia con *A theory of justice*, un famoso tratado sobre este tema del filósofo moral estadounidense John Rawls (1972). En el marco de la gran tradición kantiana, Rawls intentó circunavegar las emociones: "...por razones tanto de simplicidad como de teoría moral, he asumido la ausencia de envidia". ¿Desde cuándo podemos simplemente omitir las emociones en una discusión sobre el comportamiento humano? Me resulta desconcertante. La ironía es que si no hubiera envidia en el mundo, a nadie le importaría siquiera la justicia, dado que nunca veríamos una reacción significativa ante su ausencia. Rawls puede llamar

a la envidia un 'vicio', pero todo su tratado no tiene sentido sin ella. La enorme inversión emocional que hacen los humanos para rectificar la desigualdad y la injusticia (las protestas a gritos, las marchas, la violencia, la resistencia a las palizas policiales) son un contundente recordatorio de que no estamos lidiando con una fría construcción mental aporreada. La injusticia nos sacude hasta la médula, algo que ningún razonamiento abstracto, como el de Rawls, podrá lograr jamás.

Y así seguí interactuando con la escritura filosófica, ya sea tomando prestados conceptos o peleando con ideas que sentía que estaban divorciadas de la realidad. La última excursión de este tipo fue mi escrito sobre la 'normatividad' (de Waal, 2014). Existe una tendencia a afirmar que, mientras que el comportamiento humano está guiado por normas y expectativas sociales, el comportamiento animal simplemente es como es. Se trata de la vieja distinción humeana entre 'es' y 'debe'. No existe ningún acuerdo sobre este tema, por lo que sigue siendo un tópico de debate filosófico perenne. Sin embargo, algunos han llegado lo suficientemente lejos como para empuñar la "guillotina de Hume" con el fin de acabar con todos y cada uno de los intentos de aplicar la lógica evolutiva o la neurociencia a la moral humana. Se argumenta, incluso, que la ciencia no puede ayudarnos a comprender la moral humana (Black, 1970).

Por el contrario, creo que, si los monos capuchinos rechazan una recompensa que normalmente comen con gusto, tan sólo porque su vecino recibe una recompensa mucho mejor por la misma tarea, estamos tratando con la normatividad. El mono tiene una idea de cómo deberían distribuirse las recompensas. Los chimpancés van incluso más allá: pueden rechazar una recompensa preferida hasta que su vecino también reciba una. Hemos descubierto que los chimpancés juegan el Juego del Ultimátum de la misma manera que los niños, incluida una reacción negativa ante resultados injustos. Mientras que un niño puede gritar "¡Eso es injusto!", un chimpancé puede escupir agua en la cara de quien le hace una propuesta inaceptable (Proctor *et al.*, 2013).

Hay otros ejemplos de normatividad social en el comportamiento animal. Si adoptamos una visión más amplia, la normatividad está en todas partes. Si muevo una tela de araña, por ejemplo, la araña inmediatamente se pone a reparar su tela, lo que significa que posee un modelo de cómo debería verse una telaraña. Esto también cuenta como normatividad, incluso si no tiene que ver con el comportamiento social. Tengo la sensación de que el tema de la normatividad natural merece una exploración mucho más extensa que la que proporcioné.

El presente volumen da un gran paso al conectar la filosofía con lo que hemos aprendido sobre el comportamiento de nuestros parientes más

cercanos. En lugar de un biólogo inspirado por la filosofía, la dirección del contacto es ahora al revés. Esto seguramente conducirá a más diálogo. Noto, por ejemplo, que no todos los autores están de acuerdo con mis puntos de vista y tratan de salvar la antigua distinción entre humanos y animales que he tratado de socavar. Un volumen como este también estimulará a los filósofos a prestar más atención al reciente aumento de los estudios de la cognición animal. Este aumento no se limita a los primates, sino que incluye todo tipo de especies, también invertebrados. La sensibilidad de todos los animales está ahora en debate, y la interacción entre biólogos y filósofos será crucial para resolver esta cuestión (p. ej., de Waal y Andrews, 2022).

Felicito a los editores por haber elaborado un volumen como este y me siento honrado de que le den un lugar central al trabajo de toda mi vida.

## REFERENCIAS

- Black, M. (1970). *Margins of Precision: Essays in Logic and Language*. Ithaca, New York: Cornell University Press.
- Brosnan, S. F. & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, 425, 297-299. <https://doi.org/10.1038/nature01963>
- de Waal, F. B. M. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674033177>
- de Waal, F. B. M. (2007 [1982]). *Chimpanzee Politics: Power and Sex among Apes*. Baltimore, MD: Johns Hopkins University Press.
- de Waal, F. B. M. (2014). Natural normativity: The “is” and “ought” of animal behavior. *Behaviour*, 15, 185-204. <https://doi.org/10.1163/1568539X-00003146>
- de Waal, F. B. M. & Andrews, K. (2022). The question of animal emotions. *Science*, 375, 1351-1352. <https://doi.org/10.1126/science.abo2378>
- Midgley, M. (1979). *Beast and Man: The Roots of Human Nature*. London: Routledge. <https://doi.org/10.1038/281336a0>
- Proctor, D., Williamson, R. A., de Waal, F. B. M., & Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proceedings of the National Academy of Sciences USA*, 110, 2070-2075. <https://doi.org/10.1073/pnas.1220806110>
- Rawls, J. (1972). *A Theory of Justice*. Oxford: Oxford University Press.
- Richards, R. (1987). *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago: University of Chicago.

eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.30898>

## INTRODUCCIÓN SECCIÓN MONOGRÁFICA: “PRIMATOLOGÍA FILOSÓFICA: REFLEXIONES EN TORNO A LA OBRA DE FRANS DE WAAL”

### *Introduction to the Monographic Section: “Philosophical Primatology: Reflections on the Work of Frans de Waal”*

Joaquín SUÁREZ-RUIZ   
*Universidad Nacional de La Plata, Argentina*  
[ernestojoaquinsuarez@gmail.com](mailto:ernestojoaquinsuarez@gmail.com)

Rodrigo LÓPEZ-ORELLANA   
*Instituto de Filosofía, Universidad de Valparaíso, Chile*  
*Instituto de Estudios de la Ciencia y la Tecnología (iECYT), Universidad de Salamanca, España*  
[rodrigo.lopez@uv.cl](mailto:rodrigo.lopez@uv.cl); [rodrigo.lopez@usal.es](mailto:rodrigo.lopez@usal.es)

Al hablar de ‘antropocentrismo’ es posible distinguir, por lo menos, entre tres tipos (Mylius, 2018): un ‘antropocentrismo perceptual’, un ‘antropocentrismo normativo’ y un ‘antropocentrismo descriptivo’. El primero, el ‘antropocentrismo perceptual’, es de algún modo inevitable, por el simple hecho de que ninguna teorización puede basarse en otra cosa que no sean, en primera instancia, los datos sensoriales que un ser humano ha recibido a través de sus órganos sensoriales, ya sea directa o indirectamente. El segundo, el ‘antropocentrismo normativo’, defiende

de forma activa que existen cualidades que hacen a la especie humana más valiosa que otras especies. Aunque hay algunos investigadores que se comprometen con un enfoque de este tipo (un ejemplo de ello es la defensa de la experimentación animal por parte de Carl Cohen [1990]), se trata de un antropocentrismo cuya literatura escasea, al menos en lo referido a desarrollos académicos laicos. Probablemente, esto se deba a que sostener algún tipo de superioridad humana (ontológica, ética, etc.) no sólo podría contradecir los postulados básicos de la teoría de la evolución biológica, ampliamente aceptada por las ciencias contemporáneas, sino que, además, se comprometería con la apelación a un discutible fundamento sobrenatural y/o teológico. El tercero, el 'antropocentrismo descriptivo', que podría definirse como un 'antropocentrismo normativo pasivo', no defiende abiertamente la superioridad humana pero, mediante omisiones o distinciones no del todo explícitas, da por supuesto que la vida humana vale más o, incluso, es superior a la de todos los otros los seres vivos. Consideramos que este es el tipo de antropocentrismo más problemático, justamente, por su condición implícita.

La pandemia de COVID-19 ha explicitado que el antropocentrismo posee implicaciones concretas, no sólo para los animales no humanos y el medioambiente, sino para los mismos seres humanos. Según se ha investigado desde principios del 2020, la COVID-19 se comprende como una 'zoonosis', esto es, una enfermedad transmitida de un animal no humano a un ser humano. Esta transmisión fue favorecida por un circuito de caza y comercialización de animales silvestres que, de no haber existido, las condiciones propicias para la emergencia del SARS-CoV-2 no habrían ocurrido. Este tipo de circuitos no son exclusivos del mercado húmedo de Wuhan (China) en el cual se hipotetiza que comenzó la pandemia, sino que se extienden a lo largo y ancho del globo.

Paralelamente a las discusiones actuales en torno al antropocentrismo y sus implicaciones a nivel inter-específico, actualmente se requiere comprender en profundidad, tanto filosófica como científicamente y desde una perspectiva post antropocéntrica, la continuidad evolutiva entre los humanos y las demás especies, particularmente, entre los humanos y sus parientes filogenéticos más próximos: los grandes simios. El notable incremento de las producciones académicas sobre ética evolutiva, epistemología evolutiva o estética evolutiva, evidencian la relevancia que las investigaciones en psicología experimental o primatología han ganado en su articulación con el análisis filosófico. En esta clave, es posible argumentar sobre la relevancia de dar lugar, también, a una 'primatología filosófica', en tanto en cuanto una subdisciplina de la filosofía que pone el

foco en la continuidad evolutiva y se distingue de las versiones más antropocéntricas de la 'antropología filosófica'.

Un investigador ineludible a la hora de profundizar tanto en la articulación entre filosofía y ciencias biológicas como en una crítica científicamente informada al antropocentrismo es el primatólogo holandés Frans de Waal. Libros como *Chimpanzee politics: Power and sex among apes* (1982), *Good natured* (1996) y *Primates and Philosophers* (2006) se han convertido en clásicos cada vez más citados y discutidos entre filósofos y filósofos. Su obra permite vincular y contrastar antiguos problemas filosóficos con conocimientos diversos provenientes de la etología, las ciencias cognitivas y la biología evolutiva en general. Esta sección monográfica de *ArtefaCToS. Revista de Estudios sobre la Ciencia y la Tecnología* busca celebrar su obra y favorecer un espacio para la publicación de artículos originales que dialogan con ella.

Para aquellos filósofos/as interesados/as aún no iniciados en sus investigaciones, podemos mencionar tres nociones de particular relevancia filosófica que, de hecho, están presentes en los artículos que componen este número especial: el concepto de 'antroponegación' (*anthropodenial*), el de 'teoría de la capa' (*veneer theory*) y el de 'normatividad natural'. Respecto del último concepto, en particular, hacia el final de esta sección monográfica hemos incluido una traducción autorizada de *Natural normativity: The 'is' and 'ought' of animal behavior* (2014). Sobre las diferencias culturales que pueden influir en los diferentes presupuestos y estrategias para investigar el comportamiento de los primates no humanos, también hemos incluido la traducción autorizada del artículo *Silent invasion: Imanishi's primatology and cultural bias in science* (2003).

El número en cuestión estaba a punto de ser publicado cuando nos llegó la triste noticia del fallecimiento de Frans. Desde que comenzamos a organizar y coordinar este número especial en 2021, el primatólogo estuvo atento a su desarrollo, tanto en relación con el título elegido como también con la idea de centrarlo en la articulación entre primatología y filosofía. Ha leído los resúmenes de los artículos que lo componen y, a su vez, nos ha autorizado y conseguido el permiso para incluir las dos traducciones mencionadas. Vale decir que se trata, como detallamos en el párrafo anterior, de dos artículos de particular relevancia filosófica. A su vez, tenemos la suerte de contar con una introducción que ha realizado especialmente para el número, la cual también hemos traducido para facilitar su acceso a investigadores/as hispanohablantes. Estamos enormemente agradecidos tanto a él como a todos/as los/as autores/as que formaron parte de este proyecto. Esperamos que este número,

en homenaje a la memoria del primatólogo, contribuya a un diálogo a largo plazo entre filosofía y primatología, para el florecimiento de una investigación filosófica libre (o al menos problematizadora) de antropocentrismos.

## REFERENCIAS

- Cohen, C. (1990). Animal experimentation defended. En S. Garattini y D. W. Bekkum (Eds.), *The importance of animal experimentation for safety and biomedical research* (pp. 7-16). Cham: Springer. [https://doi.org/10.1007/978-94-009-1904-4\\_2](https://doi.org/10.1007/978-94-009-1904-4_2)
- de Waal, F. B. (1982). *Chimpanzee politics: Power and sex among apes*. New York: Harper & Row.
- de Waal, F. B. (1996). *Good natured*. Cambridge, Massachusetts.
- de Waal, F. B. (2003). Silent invasion: Imanishi's primatology and cultural bias in science. *Animal Cognition*, 6, 293-299. <https://doi.org/10.1007/s10071-003-0197-4>
- de Waal, F. B. (2006). *Primates and philosophers: How morality evolved*. Princeton, New Jersey: Princeton University Press. <https://doi.org/10.1515/9781400830336>
- de Waal, F. B. (2014). Natural normativity: The 'is' and 'ought' of animal behavior. *Behaviour*, 151(2-3), 185-204. <https://doi.org/10.1163/1568539X-00003146>
- Mylius, B. (2018). Three types of anthropocentrism. *Environmental Philosophy*, 15(2), 159-194. <https://doi.org/10.5840/envirophil20184564>




eISSN: 1989-3612

DOI: <https://doi.org/10.14201/art2024.31491>

## NORMATIVITY IN THE WILD. INSIGHTS FROM FRANS DE WAAL

### *Normatividad en estado salvaje. Aportes de Frans de Waal*

Laurence KAUFMANN   
University of Lausanne, Switzerland  
Laurence.Kaufmann@unil.ch

Recibido: 22/06/2023

Revisado: 11/08/2023

Aceptado: 11/08/2023

**ABSTRACT:** According to Frans de Waal, both humans and non-human primates possess innate instincts, emotions, and predispositions that facilitate social living. Social activities, such as forming relationships, participating in shared goals, and displaying empathy towards others, are not externally imposed obligations; rather, they are inherent and desirable aspects of social life. Against the utilitarian model of self-interest and Machiavellian intelligence, de Waal suggests a kind of gestalt reversal: far from being a mere means to achieve individual goals, social interactions are a valuable end in itself.

*Keywords:* social naturalism, moral normativity, evolution of morality, non-human primates, Frans de Waal.

**RESUMEN:** Según Frans de Waal, tanto los primates humanos como los primates no humanos poseen instintos, emociones y predisposiciones innatas que facilitan la vida social. Las actividades sociales, como entablar relaciones, participar en objetivos compartidos y mostrar empatía hacia los demás, no son obligaciones impuestas desde el exterior; más bien, son

aspectos inherentes y deseables de la vida social. Contra el modelo utilitario del interés propio y la inteligencia maquiavélica, de Waal sugiere una especie de reversión de la gestalt: lejos de ser un mero medio para lograr objetivos individuales, las interacciones sociales son un fin valioso en sí mismas.

*Palabras clave:* naturalismo social, normatividad moral, evolución de la moral, primates no humanos, Frans de Waal.

## 1. INTRODUCTION: TOWARDS A BOTTOM-UP PERSPECTIVE ON HUMAN CAPACITIES

How is society possible? To address this issue, most scholars have underscored the significant contribution of cultural institutions (e.g, church, school), shared symbolic systems (e.g, language, flags), and conventional rules (e.g, laws, customs) in facilitating socialization and enculturation. The external and regulatory influence exerted by culture is deemed crucial to counteract the asocial or even antisocial tendencies of humans.

It is this view, which assumes that social norms and rules need to suppress anarchic impulses for the collective benefit of society, which is challenged by Frans de Waal's pioneering research on the social behavior of non-human primates. According to him, both humans and non-human primates possess innate instincts, emotions, and predispositions that facilitate social living. Social activities, such as forming relationships, participating in shared goals, and displaying empathy towards others, are not externally imposed obligations; rather, they are inherent and desirable aspects of social life. Against the utilitarian model of self-interest and Machiavellian intelligence, de Waal suggests a kind of gestalt reversal: far from being a mere means to achieve individual goals, social interactions are a valuable end in itself.

De Waal's "Valuable Relationship Hypothesis" thus posits that primates and other animals are inclined to maintain relationships that hold significance to them. They actively strive to restore harmony or at least equilibrium within their relationships by engaging in actions such as resource sharing, reconciliation after conflicts or protesting unequal distributions. These actions serve as corrective measures to rectify deviations from the ideal state of a given social relationship, whether it be exchange, play, dominance, competition, cooperation, protection, or membership. To de Waal (2014), such repair work indicates a form of *primitive normativity*,

on the condition that one defines normativity as the adherence to the ideal state of a given physical or social structure. Just as a spider restores its damaged web to its original state, social animals engage in efforts to restore endangered social relationships. Both reparative activities demonstrate that non-humans can also be “guided by a template of how the structure ought to look” (de Waal, 2014, 187).

The spider analogy suggests a parsimonious and “bottom-up perspective” on normativity that is the opposite of the prevailing “top-down perspective” favored by most psychologists and philosophers (de Waal & Ferrari, 2010). By focusing almost exclusively on mental abilities such as shared intentionality, theory of mind, and symbolic capacities, the top-down perspective defines non-human species based on their cognitive limitations rather than their social and emotional capacities. Instead of obsessively searching for “human uniqueness”, de Waal’s bottom-up perspective hypothesizes that the basic capacities of our closest relatives are also present in humans. From a phylogenetic standpoint, indeed, it is reasonable to assume that nature does not discard or replace ancestral good tricks. On the contrary, our cognitive architecture functions most likely as a “Russian doll”, with its foundational, well preserved, ape-like layers serving as the basis for higher levels of information processing (de Waal, 2007). Within this cumulative evolutionary framework, ethology and primatology appear more crucial than ever to understand the human conduct. They shed light on the *common principles* that underlie both human and non-human social life, including the normative principles that I would like to discuss in the following pages.

## 2. RULES IN THE ANIMAL KINGDOM

According to de Waal and colleagues, relationship regulation in primates and chimpanzees can be accounted for in terms of two kinds of rules, “descriptive” and “prescriptive” (Flack *et al.*, 2004). “Descriptive rules” refer to statistical regularities or typical responses exhibited in specific social situations. For instance, female primates with young offspring may respond to conspecific threats by either withdrawing or displaying aggression. On the other hand, “prescriptive rules” govern male sexual competition, food resources and juvenile care and involve a reinforced sense of expected behavior from others as well as an anticipation of the consequences of deviating from these rules. This distinction overlaps with that made by social scientists. Whereas *regularities* pertain to factual observations (i.e., what *is* the case) and respond to a mere logic of repetition and

conformity, *rules* exhibit a higher level of constraint and external imposition and are reinforced through negative or positive sanctions (i.e., what *ought* to be the case).

In my view, however, de Waal's research might help to distinguish not two but three types of normativity. His "Valuable Relationship Hypothesis" suggests an intermediary level of normativity, situated midway between factual social regularities and normative cultural norms: the level of *constitutive rules*.

To support this argument, let us take the fascinating example of play regulation. As shown by Flack et al. (2004), older and younger juvenile chimpanzees engaging in play adjust their behavior, including play signals, to prevent interference from nearby adult individuals who may mistake play noises for aggressive encounters. The juvenile chimpanzees thus modify their signals in advance to prevent protective outsiders from terminating their play. The mastery of the expected intensity and vocalizations associated with play might hinge on a type of rule which is not regulative but, as will be seen, constitutive.

### 3. THE CONSTITUTIVE RULES HYPOTHESIS

Philosophers used to define regulative rules as prescribing the ways to perform pre-existing activities, such as wearing a tie for a cocktail party. These rules refer to conventional norms whose external authority is mostly enforced through informal sanctions and collective agreement. In contrast, constitutive rules establish and enable new forms of cultural activities. For example, the rules of chess not only provide instructions on how to play the game but create the very possibility of playing it. Engaging in the game of chess thus entails enacting its constitutive rules, such as having a playing partner, moving the bishop diagonally, or ending the game when checkmated. Interestingly, the violation of those rules has different consequences. Breaking regulative rules, such as wearing pajamas for a cocktail party or sacrificing all your pawns in a chess game, can be blamed on but it does not jeopardize the successful completion of the activity. On the other hand, the chess player who moves a pawn backward or puts the lost pieces back into the game dismantles the whole course of action. For he is not playing chess; he is playing *another game*.

Traditionally, constitutive rules have been associated with the creation, *via* shared intentionality, of cultural forms of activities, roles, and relationships (e.g., marriage, economic exchange, political election, etc.). Thanks to symbolic "counting-as" mechanisms, humans assign a new function to

natural or social entities, such as “this river counts as a border” or “this man counts as a president” (Searle, 1995; Tomasello & Rakoczy, 2007). Although constitutive rules are at the center of cultural creations and institutional facts, they might also be present, as de Waal’s work suggests it, in the basic repertoire of social relationships that the evolutionary pressures exerted by social life have selected. Thus, by definition, play restricts actions that may inflict harm; should a play situation become harmful, it ceases to be seen as a play and becomes a fight. Similarly, hierarchical relationships can be accounted for by constitutive rules that delineate permissible, prohibited, and obligatory behaviors based on an individual’s rank. For instance, the dominant individual possesses the right to access food ahead of others, while subordinates are obligated to wait. In the same vein, the rule of reciprocity constitutes social exchange: individual A is expected to reciprocate after receiving grooming from individual B, often in the form of providing something of value, such as food.

#### 4. NORMATIVE PARSIMONY

In accordance with the “bottom-up perspective” emphasized by de Waal, the “Constitutive Rules Hypothesis” presented here has the advantage of being ontologically and cognitively parsimonious. Indeed, the prescriptive force of constitutive rules is more appropriately characterized as a (socio-) logical “must” rather than a normative “ought to”. This logical “must” operates through *if-then* inferences, such as “if you engage in play, then you must refrain from biting” or “if you receive grooming, then you must reciprocate”. However, this *if-then* mechanism is not behaviorist in nature; it cannot be reduced to stimulus-driven, short-term associations like “if there is a nearby noise, then you flee”. Instead, the *if-then* dynamics underlying social interactions involves complex “built-in contingencies” and tree-like configurations that facilitate social problem-solving (de Waal, 2003). Coalition politics, conflict resolution, and cooperative arrangements are made feasible through a chain of social transactions whose “syntactical structure” is *conditional*: specific actions *x* must be undertaken before subsequent actions *y* can occur. Such conditional *if-then* is akin, de Waal says, to the ordering of symbols in human language: the actions within a social chain are subject to constraints comparable to the limitations on word usage within a sentence. Just as one cannot say “I dog the walk”, a capuchin monkey cannot expect much cooperation from a partner with whom he has repeatedly refused to share abundant food in the past (de Waal, 2003).

Thinking of normativity in terms of constitutive rules has interesting implications. Firstly, this sheds light on *rule-governed* behaviors, often overlooked in favor of *rule-following* actions that are much more cognitively demanding. Unlike rule-following actions, a rule-governed behavior does not require agents to have an explicit representation of a given rule and to take it as their reason for acting. Secondly, constitutive rules bridge the gap between the factual “is” of social regularities and the normative “ought” of cultural norms without resorting to an instituting scene of agreement-making and linguistic creation. They do not require the symbolic leap from nature to culture that language and mind reading are said to trigger. Thirdly, the cognitive counterpart of the normative demands on social relationships exerted by constitutive rules is not necessarily the metarepresentational ability to hold in mind simultaneously what individuals, oneself included, are doing and what they ought to do (Byrne & Whiten, 1988). As we have hypothesized elsewhere, normative demands can appear and be recognized as a succession of *affordances*, each social action affording a set of subsequent actions (Kaufmann & Clément, 2014). In contrast with the standard affordance theory, we have contended that interactions do not only offer opportunities for action; they constrain the range of appropriate responses. In other words, they not only indicate “what I *can* do”, but also “what I *must* do”. For instance, a facial expression of intimidation affords the action of complying, a request gesture affords food sharing, and a threatening face affords the action of fleeing. In short, social affordances are normative or deontic affordances that relate *in situ* perception and normativity. They allow individuals, including those who are deprived of complex metarepresentational capacities (e.g., infants, non-human primates), to see what is the right, appropriate way to respond to such and such behavior without resorting to explicit norms.

## 5. THE SOCIAL AND THE MORAL

Once the “Valuable Relationship Hypothesis” and the “Constitutive Rules Hypothesis” have been articulated, disentangling social rules from moral norms becomes almost impossible. Since primitive normativity manifests as deontic affordances within relational frameworks, moral feelings are intertwined with social dynamics. Unlike the abstract inquiry of “What should I do?” that human-specific morality has emphasized, normativity in the wild can only address the practical question of “What should I do in this specific situation and within this particular relationship?” This relationship-based socio-morality does not classify

behaviors as intrinsically good or bad, just or unjust, desirable or undesirable. Instead, it involves assessing, *within a given relationship*, who bears obligations towards whom, who possesses the authority to impose obligations, who has the entitlement to claim certain goods, and what forms of retaliation are deemed appropriate for violations of these obligations. Such relational evaluations are aptly exemplified by the duty of care, both social and moral, that is associated with the maternal bond. For instance, when the distress call of a juvenile vervet monkey is played through a concealed loudspeaker in the presence of three adult females, their gaze turns towards the mother, indicating their perception of the juvenile's call as an indication of caring behavior from the mother and their expectation for her to respond accordingly (Cheney & Seyfarth, 1990).

Since relationship-based socio-morality is sensitive to the sensorimotor affordances and requirements of close relationships, it is restricted to close relatives, familiar conspecifics and in-group members. Grounded in firsthand experiences with others, this selective socio-morality can be labeled as a "morality by perception" or "by acquaintance": it encompasses the recognition of similarity and the emotional connection with others perceived as "like-me" or "like-us" (Kaufmann, 2019). Only through abstract and sophisticated judgments can a "morality by imagination" emerge, transcending the limited criteria of sameness and proximity, and fostering concern for unknown individuals.

This form of abstract morality is most likely specific to humans, enabling the capacity to feel compassion towards distant strangers and even towards the entirety of the human species. Nevertheless, such a manifestation of abstract morality is rare. In the human world, morality by perception also tends to prevail: outgroup members do not trigger moral reactions such as empathy in the same way that in-group members do (Mondillon *et al.*, 2007).

## 6. CONCLUSIONS: TOWARDS A PHENOMENOLOGY OF NORMS

Normative or deontic affordances instantiate the constitutive rules that confer syntactic structure to social relations. However, this arid and skeletal definition does not do justice to Waal's interest in the emotional intensity of primate social life. The ability to perceive the normative affordances of a given relationship also depends on sensitivity and feeling, as evidenced by the anger that arises from their violation. For example, chimpanzees express their dissatisfaction through screaming, pounding the ground, and engaging in protest behavior when their relational

expectations are violated, such as when they are not helped by their allies or when a youngster is mistreated by an adult (de Waal, 1991). These reactions are dependent on the nature of their relationship, which pre-determines not only how they must interact, but also how they should feel about each other. For non-human but also human primates, relational feelings fluctuate according to the distance between them. When they are in a close relationship, they are expected to show empathy and compassion to each other. On the other hand, when they are socially or spatially distant, they can succumb to a state of emotional indifference, insensitivity, or even aggression. To me, such social-moral syntax, supported by a phenomenology of rules-in-action, is one of the valuable insights into the social world that can be drawn from Waal's remarkable research.

## REFERENCES

- Byrne, R. W. & Whiten, A. (Eds.) (1988). *Machiavellian intelligence. Social expertise and the evolution of intellect in Monkeys, Apes, and Humans*. Oxford: Clarendon Press.
- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press. <https://doi.org/10.7208/chicago/9780226218526.001.0001>
- de Waal, F. B. M. (1991). The chimpanzee's sense of social regularity and its relation to the human sense of justice. *American Behavioral Scientist*, 34, 335-349. <https://doi.org/10.1177/0002764291034003005>
- de Waal, F. B. M. (2003). Social syntax: the if-then structure of social problem solving. In F. B. de Waal & P. Tyack (Eds.), *Animal social complexity: intelligence, culture and individualised societies* (pp. 230-248). Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/harvard.9780674419131.c17>
- de Waal, F. B. M. (2007). The 'Russian doll' model of empathy and imitation. *Advances in Consciousness Research*, 68, 49-72. <https://doi.org/10.1075/aicr.68.06waa>
- de Waal, F. B. M. (2014). Natural normativity: The 'is' and 'ought' of animal behaviour. *Behaviour*, 151, 185-204. <https://doi.org/10.1163/1568539X-00003146>
- de Waal, F. B. M. & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, 14(5), 201-207. <https://doi.org/10.1016/j.tics.2010.03.003>
- Flack, J., Jeannotte, L., & de Waal, F. B. M. (2004). Play signaling and the perception of social rules by juvenile chimpanzees (Pan troglodytes).



- Journal of Comparative Psychology*, 118(2), 149-159. <https://doi.org/10.1037/0735-7036.118.2.149>
- Kaufmann, L. (2019). La Norme du « semblable »: entre moralité, socialité et politique. In I.Rivoal & M.Heintz (Dir.), *Morale et cognition : À l'épreuve du terrain* (pp. 73-99). Nanterre : Presses universitaires de Paris Nanterre. <https://doi.org/10.4000/books.pupo.15350>
- Kaufmann, L. & Clément, F. (2014). Wired for Society: Cognizing pathways to Society and Culture. *Topoi*, 33(2), 459-475. <https://doi.org/10.1007/s11245-014-9236-9>
- Mondillon, L., Niedenthal, G., Paula, M., & Droit-Volet, S. (2007). Imitation of in-group versus out-group members' facial expressions of anger: A test with a time perception task. *Social Neuroscience*, 2(3-4), 223-237. <https://doi.org/10.1080/17470910701376894>
- Searle, J. (1995). *The Construction of Social Reality*. New York: Simon and Schuster.
- Tomasello, M. & Rakoczy, H. (2007). What Makes Human Cognition Unique? From Individual to Shared to Collective Intentionality. In F. Clément & L. Kaufmann (Eds.), *Culture and Society: Some Viewpoints of Cognitive Scientists*. *Intellectica*, 46-47, 25-48. <https://doi.org/10.3406/intel.2007.1276>



eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.30784>

# ¿NORMATIVIDAD PRIMITIVA EN PRIMATES NO HUMANOS?

## *Primitive Normativity in Non-Human Primates?*

Laura DANÓN   
*Universidad Nacional de Córdoba, CONICET, Argentina*  
ldanon@unc.edu.ar

Recibido: 1/11/2022    Revisado: 17/04/2023    Aceptado: 9/11/2023

**RESUMEN:** Para una extensa tradición filosófica, la normatividad es un rasgo exclusivo de la especie humana. Recientemente, sin embargo, algunos filósofos y científicos comenzaron a explorar la posibilidad de atribuir algún tipo de normatividad a otras especies. Frans de Waal se destaca, en este contexto, por haber proporcionado un amplio repertorio de evidencia empírica sobre comportamientos de primates no humanos que parecen ajustarse a distintos tipos de normas. Los escépticos sobre la normatividad animal suelen, sin embargo, cuestionar este tipo de evidencia, brindando explicaciones alternativas, no normativas, de ella. Un modo en que los escépticos podrían justificar esta estrategia es apelando al llamado Canon de Morgan y aduciendo que las explicaciones no normativas introducen procesos psicológicos más simples que las normativas. Ahora bien, cuán atractiva resulte esta línea argumentativa dependerá de cómo se entienda la sensibilidad normativa. Si en lugar de centrarnos en las caracterizaciones más demandantes de tal sensibilidad adoptamos la hipótesis de que algunos primates no humanos cuentan con una suerte de "normatividad primitiva" (Ginsborg, 2011; 2018), podremos elaborar explicaciones de (al menos parte de) la evidencia empírica proporcionada por

de Waal, que rivalizan con las que ofrecen los escépticos en la simplicidad de los procesos cognitivos que invocan.

*Palabras clave:* cognición normativa, normas sociales, primates no-humanos, sensibilidad normativa, Frans de Waal.

**ABSTRACT:** According to a long philosophical tradition, normativity is an exclusive feature of human animals. However, in the last decades, many philosophers and scientists have begun to explore the possibility of attributing some kind of normativity to other species. Frans de Waal stands out among them for providing a rich repertoire of empirical evidence suggesting that non-human primates behave normatively. Skeptics regarding animal normativity usually question this kind of evidence by offering non-normative explanations of it. Here I suggest that skeptics may justify this preference for non-normative explanations by claiming that they invoke simpler cognitive processes than those who posit some sensitivity to norms (as demanded by Morgan's Canon). However, I will argue that how compelling this strategy turns out to be will depend on how one understands normative sensitivity. More specifically, I will suggest that we should abandon the most demanding characterizations of normative sensitivity in the literature and adopt the hypothesis that non-human primates may have some kind of "primitive normativity" (Ginsborg, 2011; 2018). Doing this will allow us to elaborate explanations of (some) empirical evidence provided by de Waal that may end up being as simple as those elaborated by the sceptics.

*Keywords:* normative cognition, social norms, non-human primates, normative sensitivity, Frans de Waal.

## 1. INTRODUCCIÓN

Una de las pocas tesis con respecto a la cual parece haber amplio acuerdo filosófico es aquella que afirma que los humanos somos "animales normativos". Muy sucintamente, quienes suscriben a esta afirmación suponen que los humanos somos sensibles a distintos tipos de normas – sociales, técnicas, morales, epistémicas, estéticas, etc. –, que guían nuestras acciones y funcionan como estándares a partir de los cuales evaluamos los comportamientos propios y ajenos. A lo cual suelen añadir que los animales humanos poseemos una tendencia ubicua a tomar medidas que refuerzan el cumplimiento de tales normas, premiando a quienes se ajustan a ellas y sancionando a quienes no lo hacen.

Ahora bien, ¿somos los humanos los únicos animales que contamos con estas capacidades normativas? Numerosos científicos y filósofos

piensan que este es el caso (Korsgaard, 2006; 2010; Schlingloff y Moore, 2017; Schmidt y Rakoczy, 2019). Sin embargo, en los últimos años, algunos psicólogos, etólogos cognitivos y filósofos han comenzado a explorar la posibilidad inversa: que pueda haber buenas razones para extender el rótulo de “criaturas normativas” más allá de la especie humana (Rudolf von Rohr *et al.*, 2010; de Waal, 2014; Lorini, 2018; Okrent, 2018; Vincent *et al.*, 2018; Kappeler *et al.*, 2019; Danón, 2019; Andrews, 2020; Fitzpatrick, 2020; Westra y Andrews, 2022).

Frans de Waal se destaca, en este contexto, por haber proporcionado un amplio repertorio de evidencia empírica que, según argumenta, nos muestra a diversos primates no humanos como animales cuyos patrones de comportamiento se ajustan a distintos tipos de normas. Incluso ha sostenido que algunos primates, como los chimpancés, cuentan con los “componentes esenciales” (*building blocks*) del comportamiento moral, al menos en variantes rudimentarias.

En su artículo “Natural normativity: The ‘is’ and ‘ought’ of animal behavior”, de Waal (2014) caracteriza a la normatividad como la “adherencia a un ideal o estándar” y procede a defender que “hay amplia evidencia” de que los animales se comportan, en el dominio social, de modos que parecen ajustarse a estándares o ideales. De hecho, muchas de sus investigaciones han contribuido de modo nodal a proporcionarnos tal evidencia. Así, por ejemplo, los resultados de muchas de sus observaciones y experimentos sugieren que los primates se conducen de maneras que llevan a la preservación de la armonía en sus grupos sociales: procuran insistentemente reconciliarse con otros después de un conflicto, protestan ante situaciones inequitativas, detienen las peleas entre otros, etc. Todas estas serían, para de Waal, conductas normativas que cumplen la función de corregir desviaciones respecto de un estado ideal de convivencia relativamente pacífica y armónica del grupo. Además, a su entender, contamos con otras observaciones y datos experimentales que indican que los primates poseen capacidades de auto-control y formas de resolución de conflictos que les permiten prevenir comportamientos disruptivos antes de que estos acontezcan.

Ahora bien, como han subrayado distintos filósofos, hay más de un modo en que se puede “adherir a un ideal o estándar” y corregir o prevenir las desviaciones respecto de este. En ocasiones, la conducta de un animal puede meramente ajustarse a una norma o estándar sin que debamos atribuirle por ello ninguna aprehensión de la norma o, siquiera, algún tipo de sensibilidad normativa. En tales casos, estaríamos ante criaturas cuyas conductas pueden ser descritas desde un punto de vista externo como comportamientos que están “bien regulados” (Ryle, 1949),

se “conforman” a una norma (Sellars, 1954) o son “consistentes” con ella (Schlingloff y Moore, 2017). Sin embargo, el que se produzca tal adecuación a la norma no implica que la criatura esté activamente regulando su comportamiento u obedeciendo normas o reglas. Tampoco supone que cuente con alguna capacidad para aprehender la norma en cuestión, pueda efectuar algún tipo de evaluación o juicio normativo sobre cómo ha de comportarse o sobre cómo deben ser las cosas, o aspire en algún sentido a conformarse a la norma. En breve: una criatura puede comportarse de modo acorde a una norma, incluso cuando padezca de una completa *ceguera normativa*. Si esto es así, su conducta deberá estar causada por otro tipo de factores no normativos.

Distintos tipos de factores pueden llevar a que una criatura que padece de ceguera normativa se comporte de un modo ajustado a ciertas normas. Las regularidades comportamentales que encontramos en varios grupos de primates pueden, por ejemplo, deberse a disposiciones innatas a dar respuestas específicas a determinados estímulos, a que están imitando, de modo irreflexivo e inconsciente, los comportamientos de algunos individuos especialmente poderosos o prestigiosos, etc. También es posible que, en ocasiones, para explicar estas regularidades debamos atribuir estados mentales intencionales a estos animales. Uno puede pensar, por ejemplo, que ciertos primates se comportan del modo A, que es consistente con la norma N, sencillamente porque hacer A les resulta placentero o conveniente. Este tipo de explicaciones nos muestran a estas criaturas como dotadas de vida mental, pero no necesariamente como desplegando un tipo de cognición específicamente normativa.<sup>1</sup>

En contraposición, el comportamiento de una criatura puede ser acorde con una norma porque ella se ajusta activamente a la misma. Este ajuste activo puede, a su vez, ser el resultado de una aprehensión explícita de la norma y de su decisión de comportarse de cierto modo *porque* la norma así lo prescribe. Alternativamente, puede ser el fruto de alguna suerte de sensibilidad normativa primitiva que indica a la criatura qué comportamientos particulares son los apropiados, aun cuando ella carezca de una representación explícita de una norma general que prescriba cómo debe actuar.

1. Emplearé aquí la noción de *cognición normativa* para referirme a los procesos, capacidades y estados cognitivos o psicológicos que explican la propensión de ciertas criaturas – paradigmáticamente los animales humanos– a aprehender ciertas normas (o propiedades normativas), llevar a cabo evaluaciones normativas y reforzar el cumplimiento de normas por parte de terceros.

Dado que contamos con estas distintas alternativas teóricas, parece que para obtener una mejor comprensión del fenómeno de la normatividad animal debemos intentar responder a interrogantes como: ¿de qué modo hemos de entender la tesis de que ciertos animales adhieren a normas o estándares? ¿Afirma ésta sólo que los comportamientos de algunos animales se ajustan a ciertas normas de un modo meramente externo, que ha de explicarse en términos no normativos? ¿O acaso la conformidad a la que se hace referencia supone algún tipo de reconocimiento de las normas, o de sensibilidad normativa, por parte de los animales en cuestión? Finalmente, de hallarse involucrada alguna suerte de sensibilidad a las normas: ¿cómo hemos de caracterizarla? ¿Cabe pensar que los animales aprehenden explícitamente ciertas normas y deciden seguirlas de manera autónoma y reflexiva? ¿O resulta acaso más plausible pensar que lo que poseen estas criaturas es una normatividad primitiva que les permite evaluar acciones particulares en contextos concretos, sin requerir para ello de una representación explícita de la norma general?

Dar respuesta a estos interrogantes requeriría sortear múltiples obstáculos metodológicos y conceptuales cuyo tratamiento excede ampliamente los límites de este trabajo. Hacia el final de su texto, el propio de Waal añade una nota de cautela y nos advierte, respecto de la evidencia allí reseñada, que los comportamientos presuntamente ajustados a estándares o normas sociales de los primates pueden no estar de hecho anteceditos, ni guiados, por ningún tipo de juicio normativo, ni de capacidad para experimentar cómo “deben” ser ciertas cosas. Más aún, señala explícitamente que, en el caso de los animales no humanos, la “presencia de una normatividad internalizada sigue siendo altamente especulativa” (de Waal, 2014, 201).<sup>2</sup>

Aun cuando no pretendo ofrecer aquí una respuesta acabada a la pregunta sobre si los primates no humanos cuentan con algún tipo de sensibilidad normativa, sí querría dar algunos pasos iniciales en esta dirección. A tal fin, me detendré a examinar una dificultad que enfrentan quienes, apoyándose en datos empíricos como los proporcionados por de Waal, defienden tal hipótesis. Muy brevemente, esta radica en que tales filósofos

2. En textos tempranos, encontramos interpretaciones menos cautelosas sobre la sensibilidad normativa de los chimpancés. En *Good Natured*, por ejemplo, de Waal atribuye explícitamente a los chimpancés un *sentido de la regularidad social* que define como “Un conjunto de expectativas acerca del modo en que uno (u otros) deben ser tratados y de cómo deben dividirse los recursos” (1996, 95). Sin embargo, más adelante aclara que no está convencido de que las reglas existan como tales en la cabeza de aquellos animales que las refuerzan mediante el control y las sanciones sobre las conductas de otros (1996, 96).

y científicos se ven frecuentemente interpelados por una serie de voces escépticas que proponen explicaciones alternativas de cualquier evidencia empírica que se ofrezca a favor de la existencia de alguna sensibilidad normativa en animales. Más aún, estos escépticos parecen asumir que, una vez que contamos con tales explicaciones alternativas, hemos de optar por ellas sin mayor discusión al respecto.

Una pregunta obvia que se impone en este punto es la siguiente: ¿qué razones se pueden dar a favor del proceder de los escépticos? ¿De qué modo justificar su inclinación general a favorecer sin más las explicaciones que no atribuyen sensibilidad normativa a los animales por sobre las que sí lo hacen? Una respuesta natural que podrían dar los escépticos es que sus explicaciones atribuyen capacidades cognitivas más simples a los animales que aquellas que invocan algún tipo de cognición normativa, ajustándose de este modo a un principio metodológico bien establecido en las discusiones sobre cognición animal: el Canon de Morgan. Sin embargo, según querría argumentar aquí, cuán atractiva resulte esta apelación al Canon de Morgan depende de cómo se entienda la noción de sensibilidad normativa. Según veremos, esta noción no es unívoca y hay distintos modos, más o menos demandantes, de caracterizarla. Con lo cual, aunque sea plausible sostener que las explicaciones que apelan exclusivamente a factores no normativos invocan capacidades cognitivas menos sofisticadas que aquellas que postulan una variante demandante de normatividad, es menos claro que se pueda continuar afirmando algo semejante una vez que adoptamos una noción deflacionada de sensibilidad normativa. En particular, querría mostrar que, si nos focalizamos en una porción especialmente sugerente de la evidencia empírica proporcionada por de Waal y adoptamos, como punto de partida heurístico, la hipótesis de que algunos comportamientos de los primates no humanos están guiados por una suerte de “normatividad primitiva” (Ginsborg, 2011; 2018), podremos elaborar explicaciones que rivalizan con las que ofrecen los escépticos en la simplicidad de los procesos cognitivos que introducen.

Dado el carácter incipiente de estas discusiones, sigue siendo necesario contar con más evidencia empírica relevante al respecto antes de establecer sobre bases firmes si, efectivamente, hay primates que cuentan con alguna sensibilidad normativa, siquiera modesta. Sin embargo, espero que esta línea argumentativa permita relativizar el peso de los argumentos escépticos en contra de la normatividad animal y fortalecer la conjetura de que alguna capacidad primitiva para evaluar acciones y situaciones en términos normativos puede extenderse más allá de nuestra especie.



## 2. FRANS DE WAAL Y LA NORMATIVIDAD SOCIAL DE LOS PRIMATES NO HUMANOS

Frans de Waal (2014) defiende, apoyándose en un cúmulo considerable de estudios y observaciones empíricas, que numerosos comportamientos de primates no humanos se ajustan a ciertos estándares o ideales normativos. Si nos centramos en el dominio de los comportamientos sociales, encontramos en su artículo ejemplos llamativos de ajuste normativo en: a) conductas vinculadas a la existencia de jerarquías sociales; b) conductas de reparación de vínculos; c) conductas de prevención de conflictos; d) respuestas a la inequidad; e) intervenciones imparciales de control por parte de terceros. Examinemos con algún detalle cada uno de estos comportamientos.

- a) *Conductas vinculadas a la existencia de jerarquías sociales*: Muchos primates viven en sociedades estructuradas según órdenes de dominancia relativamente estables. Más aún, los miembros del grupo reconocen estas diferencias jerárquicas mediante signos típicos y ritualizados de sumisión que los subordinados realizan ante sus superiores (de Waal, 1986). Según señala de Waal, las estructuras jerárquicas de los grupos sociales que hallamos entre los primates parecen venir acompañadas por un conjunto de normas con respecto a qué cuenta como comportamiento apropiado para quienes ocupan un lugar u otro en su interior, cómo deben tratar los subordinados a sus superiores, etc. Los primates subordinados tienden a respetar estas normas y los dominantes a reforzarlas mediante distintas formas de control y castigo. Los infantes del grupo, en cambio, suelen ignorar tales normas sociales y se les permite saltarlas sin consecuencia alguna. Sin embargo, cuando crecen comienzan a sufrir distintos tipos de castigos o represalias cuando no las respetan. Así, por ejemplo, los primates que ocupan lugares jerárquicos más elevados son quienes primero acceden a distintos recursos valiosos –como el agua o el alimento– junto a las crías pequeñas. Pero, a medida que se hacen mayores, la crías empiezan a ser castigadas cuando, por ejemplo, se aproximan a la comida que se encuentra en poder de los primates dominantes (de Waal, 1996; 2014).
- b) *Conductas de reparación de vínculos*: Según nos relata de Waal, aproximadamente treinta especies de primates buscan reconciliarse, minutos después de una pelea, estableciendo algún contacto físico amistoso (de Waal y Roosmalen, 1979; Aureli y de Waal, 2000; de Waal, 2000). Sus observaciones en el zoológico de Arnhem, por ejemplo, revelan que, después de una interacción agresiva, los chimpancés tienden a

aproximarse a su oponente y establecer algún contacto amistoso con él (besos, abrazos, vocalizaciones que indican sumisión, aproximaciones con una mano extendida, etc.) (de Waal y Roosmalen, 1979). Pero este no es un caso aislado entre los primates. Hay observaciones de fenómenos similares en bonobos, gorilas de la montaña, monos dorados, monos capuchinos, lémures pardos de frente roja, monos patas, monos vervet, babuinos, y una amplia variedad de macacos (de Waal, 1996).

Se ha conjeturado que estos comportamientos conciliatorios tienen la función de reparar el daño que han sufrido ciertos vínculos sociales por un conflicto reciente, permitiendo que los animales previamente involucrados en una interacción hostil vuelvan a interactuar de modo cooperativo (Cords y Aureli, 2000; de Waal, 2014). Se puede pensar, pues, que también en estos casos la conducta de los animales se ajusta a normas o estándares que los llevan a reparar los vínculos valiosos dañados.

c) *Conductas de prevención de conflictos*: Los primates y otros animales muestran, en distintos contextos de interacción, conductas cuya función parece radicar en evitar la generación de conflictos. Cuando juegan, por ejemplo, se comportan de modos semejantes a cuando se involucran en encuentros agonísticos (pegan, muerden, se dan palmadas, etc.), pero con un menor grado de agresividad. Dado que la línea entre las interacciones lúdicas y las agresivas es difusa, estas conductas se ven frecuentemente acompañadas por signos característicos que, según los investigadores, cumplen la función de indicar a los otros animales involucrados que “se trata sólo de un juego” (Palagi, 2007). A su vez, cuando los infantes juegan con otros, sus madres permanecen atentas a estas interacciones y las interrumpen si se vuelven excesivamente agresivas. Ante ello, los participantes en el juego suelen emitir vocalizaciones específicas que parecen servir para evitar tales intervenciones, reasegurando a las madres que la interacción que observan es de naturaleza benigna (Flack et al., 2004). Datos como estos sugieren que las instancias de juego entre primates pueden estar gobernadas por normas acerca de cuál es el comportamiento apropiado en contextos lúdicos, las cuales se ven reforzadas tanto por las “señales de juego” de los participantes como por las intervenciones maternas (Fitzpatrick, 2020).

d) *Respuestas a la inequidad*: Según señala de Waal, para los animales de especies altamente sociales y cooperativas resulta de relevancia adaptativa llevar algún registro de los beneficios que obtienen, por comparación a los que reciben aquellos con quienes han colaborado, a fin de no quedar reiteradamente en desventaja cuando se involucran en actividades conjuntas (de Waal, 2014). Es natural esperar, a su vez,

que manifiesten reacciones negativas cuando se encuentren ante distribuciones inequitativas de recompensas. Partiendo de esta hipótesis, Brosnan y de Waal (2003) llevaron a cabo un experimento en el cual dos monos capuchinos realizaban la misma tarea, pero recibían en algunas ocasiones la misma recompensa y, en otras, una diferente. En el último caso, quienes obtenían una recompensa menos deseable que la de su compañero reaccionaban negativamente rechazándola, o, incluso, negándose a seguir cooperando. Se han hallado, además, reacciones negativas a la inequidad en otras especies, tanto primates como no-primates, incluyendo macacos (Massen *et al.*, 2012; Hopper *et al.*, 2013), chimpancés (Brosnan *et al.*, 2005; 2010), titis cabeciblancos (Neiworth *et al.*, 2009); perros (Range *et al.*, 2009), córvidos (Wascher *et al.*, 2013) y ratas (Oberliessen *et al.*, 2016).

Un punto relevante a señalar es que, por lo general, reacciona negativamente aquel primate que ha recibido una recompensa menor o menos deseable y no quien se vio beneficiado. Llamativamente, sin embargo, hay evidencia que indica que los chimpancés no sólo reaccionan negativamente cuando reciben una recompensa menos deseable que otro por el mismo trabajo; también tienden a rechazar una recompensa más valiosa cuando es su compañero el que recibe algo menos deseable (Brosnan y de Waal, 2003; Brosnan *et al.*, 2010).

Que el primate perjudicado en el reparto de recompensas rechace el alimento, o se niegue a cooperar, acarrea una pérdida a corto plazo para él (pues, en principio, parece preferible quedarse con un alimento menos deseable a no recibir nada). Sin embargo, tal pérdida inmediata puede comportar beneficios a largo plazo. Presuntamente, en contextos naturales, un animal se ve perjudicado cuando algún compañero con quien ha cooperado acapara para sí la mejor parte de lo obtenido en una tarea conjunta. Luego, su enojo y su negativa a colaborar pueden ser modos de indicarle cuáles son sus expectativas. Esto puede, eventualmente, llevar a que el compañero se ajuste a ellas o, en último término, permitirle buscar otro par más justo para dividir los bienes (Price y Brosnan, 2012). En este sentido, cuando un primate perjudicado en el reparto de recompensas se enoja, rechaza la recompensa menor, o deja de colaborar, está comportándose de un modo que tiene por función lograr que los intercambios con otros se ajusten a ciertos parámetros de equidad. Ahora bien, aunque el enojo ante las situaciones de inequidad conduzca eventualmente a corregirlas, y esa sea su función evolutiva, esto no significa que los primates que reaccionan de este modo estén interesados en la equidad antes que, por ejemplo, en que su situación no sea desventajosa en relación a la de sus compañeros. El caso de los chimpancés

que rechazan la recompensa mayor, en cambio, podría sugerir algo diferente: una preocupación general, no sólo por la situación en la cual uno se ve perjudicado, sino por todo tipo de situación de inequidad, incluso aquellas que perjudican a otros.

- e) *Intervenciones imparciales de control por parte de terceros*: En algunas especies de primates hallamos ejemplos de comportamientos imparciales de control de la agresión por parte de terceros (*policing*). Estas son conductas pacificadoras mediante las que un primate interviene en un conflicto o situación agonística en la que no se encontraba inicialmente involucrado. Los chimpancés dominantes, por ejemplo, suelen detener las peleas aproximándose a los combatientes, con su pelambre erizada, interponiéndose entre ellos y ahuyentándolos, o separándolos con sus propias manos. Aunque se trata de conductas relativamente poco frecuentes entre los primates, estas intervenciones imparciales por parte de terceros han sido observadas en chimpancés (Boehm, 1994; von Rohr *et al.*, 2012), bonobos (Vervaecke *et al.*, 2000), gorilas de montaña (Watts *et al.*, 2000), orangutanes (Tajima y Kurotori, 2010), monos dorados (Ren *et al.*, 1991) y macacos cola de cerdo (Flack *et al.*, 2005).

Un rasgo llamativo de estas intervenciones es su imparcialidad. A diferencia de lo que ocurre cuando un familiar interviene a favor de otro durante un encuentro agonístico, los primates dominantes que efectúan estos comportamientos de control no toman partido por uno u otro de los involucrados en el conflicto, ignorando sus vínculos sociales previos e interviniendo de un modo que no refleja sus preferencias sociales usuales (de Waal, 2014). Tampoco encontramos aquí, como ocurre con las conductas de castigo, una agresión directa a quien se comporta inadecuadamente. La función que se ejerce en estos casos puede ser descripta, antes bien, como un comportamiento de “arbitraje” (Rudolf von Rohr *et al.*, 2012) que apunta a restaurar la paz, antes que a favorecer a los aliados, parientes o amigos (de Waal, 1996).

Un punto adicional a señalar es que estas conductas de arbitraje comportan cierto riesgo para quien las lleva a cabo, pues requieren que este se aproxime a los individuos involucrados en una pelea, arriesgándose a convertirse en blanco de agresiones (Flack *et al.*, 2005). Pero, además, pueden ser costosas en términos de energía y oportunidades. Por otra parte, no parecen ser conductas que apunten a conseguir favores a cambio. Pues, como señala de Waal, si este fuera el objetivo, el modo más directo de alcanzarlo sería aliarse con los más poderosos entre los

involucrados en la pelea, quienes suelen estar en mejor posición de otorgar mayores favores posteriormente (de Waal, 1996; 2014).

Se han propuesto distintas hipótesis con respecto a la función de las conductas de control o arbitraje. La conjetura más popular es que sólo traen beneficios indirectos para quien funge como árbitro, en tanto permiten aumentar la estabilidad del grupo, reducir la cantidad de conflictos y construir redes sociales más amplias y diversas (Flack *et al.*, 2005; 2006; Rudolf von Rohr *et al.*, 2012). Atendiendo, en cambio, a las causas próximas de estos comportamientos, se ha pensado que pueden estar motivados por el interés de mantener una dinámica de interacciones pacíficas, armoniosas y estables en el grupo, con lo cual podrían reflejar cierta “preocupación por la comunidad” (de Waal, 1996; Rudolf von Rohr *et al.*, 2012). Esto ha llevado a que, en ocasiones, se considere a este control social como un precursor de la moral y como un tipo de comportamiento pro-social (Rudolf von Rohr *et al.*, 2012).

Hasta aquí hemos revisado (de modo no exhaustivo) un repertorio de evidencia empírica que, para Frans de Waal, sugiere que el comportamiento de distintas especies de primates no humanos se ajusta a ciertos estándares o normas sociales o, incluso, permite corregir desviaciones, propias y ajenas, respecto de los mismos. Sin embargo, según señalamos, una larga tradición filosófica considera que el que algunos comportamientos animales sean consistentes con ciertas normas no es suficiente para sostener que quienes los llevan a cabo sean criaturas normativas en un sentido genuino del término. Para esto último hace falta, adicionalmente, que efectúen tales comportamientos en virtud de alguna suerte de sensibilidad normativa. A lo cual muchos añaden que no podemos encontrar una sensibilidad de este tipo entre los animales no humanos. El próximo apartado estará dedicado a examinar con mayor detalle tales posiciones escépticas.

### **3. LOS ESCÉPTICOS RESPECTO DE LA NORMATIVIDAD ANIMAL Y EL CANON DE MORGAN**

Aun cuando numerosos investigadores aceptan hoy que los primates no humanos son criaturas con vidas sociales complejas y capacidades cognitivas fascinantes, pocos están dispuestos a sostener que cuentan con alguna sensibilidad normativa. De hecho, distintas voces escépticas han ofrecido razones para rechazar que este sea el caso (Korsgaard, 2006; Schlingloff y Moore, 2017; Schmidt y Rakockzy, 2019). Con frecuencia, estos escépticos señalan que, en sentido estricto, la evidencia disponible

no resulta convincente a la hora de atribuir alguna sensibilidad normativa a animales no-humanos, debido a que se pueden ofrecer explicaciones alternativas de ella que no invocan factores normativos.

Ahora bien, ¿de qué índole son estas explicaciones alternativas? Un punto interesante a señalar es que, probablemente debido a la alta versatilidad y complejidad de los comportamientos sociales de los primates, a menudo estas explicaciones involucran la atribución de estados psicológicos que guían las respuestas de estos animales. Con lo cual, los escépticos con respecto a la normatividad animal suelen admitir la legitimidad de tratar a los primates como criaturas con vidas psicológicas, que actúan a la luz de lo que creen, perciben, desean, etc. (ver, por ejemplo, Korsgaard, 2006, 142-143). Lo que estos investigadores rechazan es una tesis más específica: que, entre las capacidades, procesos y estados cognitivos que atribuimos a estos animales haya algunos que les permitan aprehender ciertas normas, o, de modo más laxo, propiedades normativas.

Llegado este punto, podríamos preguntarnos por qué todos estos filósofos insisten en que las explicaciones normativas se tornan “prematuras” o “innecesarias” desde el momento en que contamos con una explicación alternativa (más o menos plausible) del comportamiento, incluso cuando esta atribuya a los animales diversos estados mentales intencionales. En otras palabras: ¿por qué se asume que las explicaciones que apelan a algún tipo de cognición normativa sólo están justificadas si no es posible dar ninguna otra explicación?

Los escépticos sobre la normatividad animal podrían responder a este interrogante apelando a un principio metodológico que ha tenido un impacto casi omnipresente en las discusiones sobre cognición animal desde que fuera formulado, a fines de siglo XIX, por Lloyd Morgan, considerado hoy por muchos como el padre de la psicología comparada: el *Canon de Morgan*. De acuerdo con la formulación del propio Morgan este principio nos dice que:

En ningún caso hemos de interpretar una acción como el resultado del ejercicio de una facultad psicológica más elevada si podemos interpretarla como el ejercicio de una capacidad que ocupa un lugar más bajo en la escala psicológica. (Morgan, 1894, 53)

Bajo la interpretación que aquí nos interesa, este canon tiene la función de guiar nuestras atribuciones de estados mentales a los animales, instándonos a elegir, entre las hipótesis explicativas disponibles, aquella que logre explicar el comportamiento animal postulando capacidades que “ocupen un lugar más bajo en la escala psicológica”. Aunque no es claro de qué modo se debe trazar la distinción entre capacidades psicológicas

de alto y bajo nivel, parece plausible pensar, en términos generales, que estas expresiones refieren a capacidades de mayor o menor grado de sofisticación cognitiva (Fitzpatrick, 2017). Se asume así, por ejemplo, que entre una explicación del comportamiento que apele a representaciones de primer orden y una que postule representaciones de segundo orden, debemos optar por la primera. O que, siquiera en primera instancia, hemos de preferir las explicaciones que apelen a procesos de aprendizaje por asociación por sobre aquellas que postulen procesos de razonamiento práctico, etc.

Como señala Fitzpatrick (2017), bajo esta interpretación del Canon, debemos rechazar las explicaciones que apelen a capacidades cognitivas más sofisticadas siempre que contemos con una alternativa que invoque capacidades más simples y pueda dar cuenta satisfactoriamente de los mismos datos. En sus variantes más extremas, esta posición nos conduce a un “negacionismo de sillón” para el cual “la mera habilidad para construir una explicación alternativa de bajo nivel de los datos empíricos resulta suficiente para sugerir que la explicación más elevada es inadecuada.” (Fitzpatrick, 2017, 441).

Es interesante reparar en que, aunque los escépticos respecto de la normatividad animal no suelen mencionar explícitamente al Canon de Morgan en sus discusiones, esta es una herramienta a la que pueden apelar para justificar su posición. Basta para ello con que asuman que la cognición normativa es más sofisticada que la no normativa y, luego, siguiendo lo prescripto por el Canon, se nieguen a atribuir a los animales no humanos cualquier tipo de sensibilidad normativa siempre que logren pergeñar una explicación alternativa de sus comportamientos.

Aunque esta estrategia puede parecer prometedora, en lo que sigue querría argumentar que pierde buena parte de su fuerza una vez que reparamos en que hay distintos modos en los que podemos entender la sensibilidad normativa en el mundo animal. Esto, obviamente, no implica que el escéptico no pueda apelar a otras estrategias argumentativas para justificar su posición. Luego, la argumentación que ofreceré en lo que sigue es modesta y apunta solo a mostrar que, al menos en algunos casos, la apelación al Canon de Morgan es insuficiente para descartar sin más la hipótesis de que algunos primates poseen cierto tipo de sensibilidad normativa. Sin embargo, este paso mínimo supone un avance para los debates sobre normatividad animal, pues nos invita a explicitar y sopesar mejor el abanico de razones que cabe invocar para rechazar el tipo de evidencia empírica proporcionada en los puntos a)-e). Con la finalidad de avanzar en esta dirección, dedicaré el próximo apartado a discutir

algunos modos, más o menos demandantes, de entender la sensibilidad a las normas.

#### 4. SENSIBILIDAD NORMATIVA: ENFOQUES INTELLECTUALISTAS VERSUS NORMATIVIDAD PRIMITIVA

Un punto relevante a considerar, cuando nos preguntamos si los primates no humanos cuentan o no con algún tipo de sensibilidad normativa, es el siguiente: en sentido estricto, hay distintos modos en los que una criatura podría ser sensible a las normas o realizar alguna suerte de evaluación normativa. En lo que sigue querría examinar brevemente dos propuestas sobre cómo entender la sensibilidad normativa: el enfoque intelectualista, defendido paradigmáticamente por Christine Korsgaard, y un enfoque deflacionado, desarrollado por Hannah Ginsborg, que gira en torno a la noción de normatividad primitiva.

Para los enfoques intelectualistas, la criatura que adhiere a una norma debe contar con un conocimiento explícito de la misma *qua norma*. Esto supone que es capaz de volverse reflexivamente sobre las normas en cuanto tales, representándose explícitamente su contenido y su condición de normas. Además, ha de poder reconocer que la norma prescribe cierta acción A y esto ha de motivarlo a realizar A (Okrent, 2018).

Ahora bien, los intelectualistas suelen requerir, además, que quien sigue normas pueda distanciarse de sus deseos, impulsos y motivaciones, para evaluar si son acordes con las normas que ellos aceptan y si, en consecuencia, deben actuar (o no) según tales inclinaciones (Korsgaard, 2006; 2010). De hecho, estos filósofos suelen subrayar que tal capacidad para el “auto-gobierno normativo” es distintiva de los animales humanos. Los animales no humanos, en cambio, son caracterizados como criaturas caprichosas que sencillamente actúan guiadas por sus deseos, emociones o impulsos más fuertes: aquellos que logren imponerse por sobre todos los otros (Korsgaard, 2006; Kitche, 2006).

Aunque muchos filósofos suscriben explícitamente a esta interpretación clásica de la sensibilidad a las normas, o al menos la asumen implícitamente, ella no agota nuestras posibilidades teóricas. En contraposición, Hannah Ginsborg (2011; 2018) postula la existencia de un tipo de “normatividad primitiva”, previa e independiente de la normatividad plena descrita por los intelectualistas, que permite a quien la posee: i) responder de un modo específico a una situación particular y ii) percatarse de que lo que está haciendo es adecuado. Sin embargo, i) y ii) no dependen del reconocimiento o aprehensión explícita de una regla o norma general. Por



lo tanto, las criaturas dotadas de este tipo de normatividad pueden darse cuenta de que están respondiendo de modo apropiado a su entorno, pero carecer del conocimiento explícito de una norma general que prescriba este tipo de comportamientos.

En sus textos, Ginsborg atribuye este tipo de normatividad primitiva a los infantes humanos pre-lingüísticos. Un niño pequeño puede, piensa, aprender a discriminar los objetos verdes de los azules, poniendo cada uno en una pila diferente, a partir de la observación de las acciones de un adulto o del entrenamiento previo. Además, puede percatarse, cuando lleva a cabo tal tarea, de la adecuación de sus respuestas y de que estos objetos “encajan” juntos. Sin embargo, dada su corta edad, este niño puede carecer de conceptos generales como los de *rojo* y *azul*, que necesitaría dominar para aprehender una regla general como: *se debe poner todos los objetos rojos en una pila y los azules en otra*. Esto no le impide, sin embargo, desplegar un tipo primitivo de sensibilidad normativa que va a consistir en acuñar ciertos contenidos mentales, como *esto es adecuado* o *esto encaja*, con los que evaluar normativamente las acciones particulares que realiza en ese momento (Ginsborg, 2011, 252).

Ahora bien, para acuñar contenidos como *esto es adecuado* o *esto encaja*, solo parece preciso referir demostrativamente a una acción actual que uno está realizando, u observando realizar a otros, en el “aquí y ahora” inmediato, y atribuir a esta acción particular una propiedad normativa como la de *ser adecuado* o *encajar*. Por otra parte, estos son contenidos de primer orden, que representan ciertos estados de cosas mundanos. Luego, una criatura puede formar este tipo de pensamientos, aunque carezca de capacidades meta-representacionales de segundo orden o, incluso, sea incapaz de representarse cómo deberán actuar ciertos agentes en el futuro, cómo debieron actuar en el pasado, cómo deberían actuar en situaciones hipotéticas, etc. En estos dos puntos, la normatividad primitiva resulta menos demandante que la capacidad para la aprehensión y seguimiento de normas de la que hablan los enfoques intelectualistas.

Parece claro que estas diferencias con respecto a las competencias cognitivas requeridas por el enfoque intelectualista y el de la normatividad primitiva incidirán en cuán plausible resulte atribuir un tipo de sensibilidad normativa u otra a los animales sin lenguaje. En primera instancia, parece complejo extender el enfoque intelectualista a los animales no humanos por dos razones. Por una parte, es dudoso que estos animales tengan la capacidad para tener pensamientos de segundo orden o capacidades reflexivas para pensar acerca de nuestros estados mentales en cuanto tales, como este enfoque requiere (Suddendorf, 1999; McGeer y Petit, 2002; Bermúdez, 2003). Por otra, resulta discutible que cuenten con un

concepto abstracto como el de “norma” e, incluso, con todo el espectro de conceptos generales (sobre tipos de acciones, circunstancias y criaturas) que se requieren para articular explícitamente el contenido de distintas normas generales. Así, por ejemplo, si las conductas de tolerancia que muestran los chimpancés hacia los infantes del grupo estuvieran guiadas por una norma en el sentido intelectualista del término, estos monos tendrían que contar con algunos conceptos como el de *infantes*, *no agredir* y *deber* (u otros semejantes). Análogamente, los primates que responden negativamente a la inequidad deberían tener conceptos como los de *deber*, *recompensa*, *equitativa*, *inequitativa*, *misma tarea*, etc. Es claro que este tipo de estrategia puede llevarnos a atribuir una red muy rica y heterogénea de conceptos a los animales de un modo que no siempre será fácil justificar.

Nada de esto se requiere, en cambio, para atribuir normatividad primitiva a un animal no humano. Según vimos, en este caso el animal solo deber contar con un concepto demostrativo que le permita referirse a una situación particular que está percibiendo de modo inmediato y con la capacidad para evaluar normativamente esta situación mediante la aplicación de algunos predicados como *apropiado* o *no apropiado*. Estamos aquí, pues, ante pensamientos de primer orden que no involucran la representación de una norma general, conceptos deónticos como el de *deber*, o el de *norma*, ni una red rica y compleja de conceptos que tipifiquen clases de acciones, individuos y situaciones a los que las normas representadas se aplican.

Pese a que, dadas todas estas diferencias, parece menos complejo atribuir normatividad primitiva a los animales no humanos que atribuirles normatividad en el sentido intelectualista, es preciso reconocer que Hannah Ginsborg se mantiene dubitativa con respecto a esta posibilidad. Por contraposición, Kristin Andrews ha reelaborado la noción de Ginsborg, dándole un tenor social, y ha defendido la posibilidad de atribuirla a los animales no humanos (Andrews, 2020). Para esta autora, podemos atribuir normatividad primitiva o ingenua (de carácter social) a algunos animales no humanos, como los primates, en la medida en que estos no sólo son capaces de distinguir a los miembros de su propio grupo de los individuos ajenos a este, sino que cuentan, además, con un sentido de *cómo hacemos las cosas nosotros* que no depende de su aprehensión previa de reglas o normas generales.

En el próximo apartado querría emplear la noción de normatividad primitiva para dar cuenta de cierta porción de la evidencia empírica proporcionada por de Waal. Según intentaré mostrar, esto nos permite ofrecer interpretaciones de esta evidencia que rivalizan con las ofrecidas

por los escépticos sobre la normatividad animal en lo que respecta a la simplicidad de los procesos, capacidades y estados cognitivos que atribuyen.

## 5. SIMPLICIDAD COGNITIVA, EVIDENCIA EMPÍRICA Y NORMATIVIDAD PRIMITIVA

Como vimos en la sección 2, quienes rechazan la atribución de algún tipo de sensibilidad normativa a los primates no humanos suelen adoptar la siguiente estrategia: objetan la evidencia comportamental a la que se apela en la literatura para sugerir la existencia de alguna capacidad de esta índole, ofreciendo explicaciones alternativas, no normativas, de la misma. Al hacerlo, parecen asumir (aunque no lo explicitan) que sus explicaciones invocan capacidades, procesos y estados cognitivos más sencillos y menos sofisticados que los que atribuyen a los animales algún tipo de sensibilidad normativa.

Ciertamente, en algunos casos, no parece difícil formular explicaciones no-normativas más simples (en términos de los estados y procesos cognitivos postulados) y plausibles que las que invocan alguna comprensión o sensibilidad a las normas. Pensemos, por ejemplo, en el primate que busca reconciliarse con un superior con el cual se ha peleado recientemente. ¿Debemos explicar su comportamiento apelando a su sensibilidad a una norma como *uno debe buscar reparar los vínculos valiosos con otros*? ¿No parece acaso más sencillo pensar que este primate *desea* reestablecer el vínculo con su superior o que *está ansioso* por haberse peleado con él? O consideremos el caso de la madre chimpancé que interviene para detener aquellos intercambios lúdicos en los cuales participa su hijo cuando comienzan a volverse demasiado agresivos. ¿Por qué pensar que está respondiendo a una norma como *es preciso interrumpir las interacciones agresivas entre infantes*? ¿No es acaso más natural pensar, en cambio, que actúa como lo hace porque percibe *que la situación es peligrosa para su hijo* y desea *evitar que sufra algún daño*?

Ahora bien, las cosas no son siempre tan sencillas. Retomemos, a modo de ejemplo, dos casos. El primero de ellos es el de las respuestas negativas de ciertos primates ante el reparto inequitativo de recompensas. En principio, el escéptico puede poner en duda que, cuando un mono capuchino o un chimpancé se enoja y se niega a seguir colaborando al recibir una recompensa menos deseable que su compañero, esté actuando en virtud de una genuina evaluación normativa sobre la injusticia de la situación. De hecho, se podría pensar, alternativamente, que lo

que está motivando la reacción de enojo del mono es que *no le gusta* recibir un pepino en lugar de una uva. Sin embargo, algunos datos adicionales nos llevan a cuestionar la adecuación de esta explicación tan simple y directa del comportamiento de estos primates.

Un primer punto que destacan Brosnan y de Waal es que las respuestas negativas a la inequidad solo tienen lugar cuando un primate observa que su compañero recibe una recompensa mayor *por la misma tarea*. El que este animal vea, en cambio, que hay una recompensa mayor, aunque nadie la reciba, no produce tales reacciones. Un segundo rasgo curioso es que el mero ser alimentado de modo no equitativo, si no se le ha exigido trabajo alguno previamente, tampoco genera respuestas negativas (Wolkenten *et al.*, 2007; Brosnan *et al.*, 2010). Estos datos empíricos resultan relevantes para nuestra discusión, pues sugieren que lo que provoca la reacción de los primates no es simplemente haber recibido un alimento en lugar de otro. Con lo cual no podemos explicar su comportamiento atribuyéndoles un deseo como *quiero la uva* y un estado de enojo porque *recibí el pepino y no la uva*. Algo más complejo parece estar teniendo lugar: estos primates reaccionan negativamente porque hay otro que recibe una mejor recompensa que ellos *después de llevar a cabo la misma tarea*.<sup>3</sup>

3. Algunos investigadores han sostenido que las respuestas negativas de estos primates no se deben a la inequidad de la situación, como conjeturan Brosnan y de Waal, sino a la frustración que les provoca el haber pasado de recibir un alimento muy deseable (como las uvas) a ser alimentados con algo menos deseable (como los pepinos) (Roma *et al.*, 2006; Dubreuil *et al.*, 2006). Si bien una interpretación de este tipo parece cognitivamente parsimoniosa, es preciso señalar algunas objeciones en su contra. En primer lugar, como indicamos arriba, las reacciones a la inequidad que registraron por Brosnan, de Waal y colegas solo se producen cuando se da a los primates un alimento después de que estos han llevado a cabo una tarea. Ahora bien, en los estudios de quienes defienden la hipótesis de la frustración, se ofrecen a los primates alimentos más o menos deseables, sin que estos tengan que llevar a cabo tarea alguna (con la excepción de Silberberg *et al.*, 2009). Luego, aunque en estos experimentos la respuesta de los primates parece no verse alterada por el que sus compañeros obtengan un alimento mejor, sino por el hecho de que ellos mismos han recibido un alimento mejor en el pasado y ahora reciben uno menos deseable, cabe argüir que esto se debe a que se está midiendo un fenómeno diferente – la frustración– que tiene lugar en circunstancias distintas que las respuestas a la inequidad. Por otra parte, aún si alguien insistiera en interpretar los resultados de Brosnan y de Waal en términos de frustración, parece preciso reconocer que, en sus experimentos, la frustración no se debería simplemente a que quería un alimento y recibió otro, sino a *que recibí un alimento menos deseable tras realizar una tarea y mi compañero recibió un alimento más deseable por la misma tarea*, o algo similar. Con lo cual, el tipo de contenidos representacionales que deberíamos atribuir a los primates para explicar sus repuestas sigue siendo notablemente complejo.

El escéptico aún puede pergeñar una explicación de la conducta animal que apele a estados de deseo y tome en cuenta los datos recién mencionados. Puede, por ejemplo, sostener que los sujetos del experimento se comportan como lo hacen porque *desean recibir lo mismo que otros por la misma tarea*. Pero, sin dudas, esta interpretación los obliga a atribuir a dichos primates un contenido más alambicado y complejo. Ante este tipo de señalamiento, el escéptico puede insistir en que, aunque los deseos que ha terminado atribuyendo a los primates tienen contenidos relativamente complejos, su explicación aún apela a capacidades psicológicas menos elevadas que aquellas que invocan alguna sensibilidad a las normas. De hecho, esta parece ser una línea argumentativa especialmente convincente si se adopta una posición intelectualista, dado lo demandantes que son los requisitos cognitivos que, según este enfoque, deben satisfacer las criaturas sensibles a normas. En el caso que nos ocupa deberíamos asumir, por ejemplo, que los primates han aprehendido explícitamente una norma como *se debe dar recompensas equivalentes por la misma tarea*, para lo cual han de contar con representaciones complejas como la de *recompensas equivalentes y misma tarea*. Pero, además, deben ser capaces de considerar a este tipo de contenido como una norma, evaluando los propios estados motivacionales a fin de dirimir si se ajustan a ella (o no) y actuando como indica. Lo cual supone que poseen el concepto de norma y capacidades para volverse reflexivamente sobre sus propios estados mentales motivacionales y tomarlos como objetos de pensamientos ulteriores.

Las cosas cambian, sin embargo, si lo que atribuimos a los primates es un tipo de normatividad primitiva. En este caso, como vimos, ya no necesitamos suponer que cuentan con competencias meta-representacionales, cognición de segundo orden, o capacidades para evaluar si sus deseos se ajustan o no a una norma. Tampoco tienen por qué poseer el concepto de norma, ni la red de conceptos generales necesarios para representar el contenido general de la misma. Solo deben poder reaccionar, ante situaciones concretas, formando contenidos como *esto es apropiado* o *esto no es apropiado*.

¿Qué otras capacidades cognitivas ha de poseer una criatura para poder formar estos contenidos? Aunque no puedo dar aquí una respuesta detallada a este interrogante, querría realizar algunos señalamientos iniciales al respecto. En primer lugar, parece preciso poder percibir situaciones particulares en el entorno inmediato. Ahora bien, la capacidad para percibir sucesos o estados de cosas del entorno parece ser de crucial relevancia adaptativa, en la medida en que la percepción nos brinda información que resulta nodal para guiar exitosamente nuestra conducta. Para

muchos investigadores, además, es una capacidad relativamente primitiva y se encuentra en numerosas especies no humanas (Burge, 2010; Barrett, 2015; Glock, 2019). Más aún, en principio, un animal podría ser capaz de percibir cómo son las cosas, o qué está pasando en su entorno inmediato, aun cuando careciese de habilidades representacionales más complejas para, por ejemplo, recordar episodios de su pasado, imaginar situaciones contrafácticas, planificar situaciones futuras, pensar acerca de sus propios pensamientos, etc.

A esto se suma, como ya señalamos anteriormente, que el enfoque de la normatividad primitiva aquí defendido solo nos exige que el animal registre y represente la situación particular a la que se enfrenta de modo inmediato, aunque no cuente con conceptos generales que le permitan clasificarla como una instancia particular de cierto tipo o clase de situación o estado de cosas. Esto permite, como señala Ginsborg (2011), que el niño que coloca un cubo azul junto a otro piense, al hacerlo, *esto es adecuado*, aunque carezca del concepto de “azul”. De modo análogo, podría permitir que un primate piense, ante una situación de injusta, algo como: *esto es inadecuado*, aunque carezca del concepto de injusticia, etc. En este sentido, estamos ante un tipo de evaluaciones normativas que no exigen el mismo tipo de repertorio conceptual robusto que se debería tener para formular normas explícitas, de carácter general, que establezcan que cierto tipo de situaciones o estados de cosas deben ser el caso.

Más controvertido resulta, en cambio, dirimir qué tipo de representaciones debe emplear una criatura para llevar a cabo estas evaluaciones normativas primitivas de estados de cosas o situaciones inmediatas. En primera instancia, parece que entre estas representaciones tendrá que haber una que funcione como un demostrativo – permitiendo singularizar o identificar el estado de cosas o situación sobre la cual ha de versar el pensamiento<sup>4</sup> y otra que debería operar como un predicado evaluativo, que caracterice a ese estado de cosas o situación como adecuado o inadecuado. Como se indicó arriba, no parece polémico afirmar que numerosas especies de animales cuentan con la capacidad para percibir situaciones o estados de cosas particulares presentes ante su experiencia sensorial. Pero, si esto es así, también resulta plausible pensar que pueden, además, singularizar las situaciones o estados de cosas percibidos y atribuirles alguna propiedad ulterior. Esto es lo que permitiría a ciertos animales formar, respecto de algunas situaciones particulares que

4. Ver Nogueira de Carvalho (2016) por una discusión cuidadosa y detallada de estos demostrativos que él autor caracteriza, de modo relevante para los intereses de este trabajo, como independientes del lenguaje.

perciben, contenidos como *esto es peligroso, esto es fácil, esto es seguro*, etc.<sup>5</sup>

El caso que aquí nos interesa específicamente – la posibilidad de aplicar algún tipo de conceptos evaluativos a situaciones o estados de cosas – no ha sido extensamente estudiado en animales no humanos. Sin embargo, hay quienes defienden que puede haber capacidades básicas para efectuar evaluaciones normativas de los sucesos que ocurren en el entorno próximo, que podrían ser de índole perceptual o afectivo-experencial (Bergqvist y Cowan, 2018). Incluso se ha sugerido que estas experiencias evaluativas podrían, en sus variantes más primitivas, estar presentes en criaturas carentes de lenguaje como los niños pequeños (Audi, 2018). Aunque no podré extenderme aquí en este punto, pienso que este tipo de comprensión minimalista de las experiencias evaluativas puede dar respaldo a quienes afirman que hay especies animales capaces de valorar normativamente ciertos modos de hacer las cosas por sobre otros (Vincent *et al.*, 2018; Glock, 2019).

Tomando en cuenta estas ideas previas, podemos genuinamente preguntarnos si al extender estas capacidades para la percepción de situaciones particulares y su evaluación normativa a algunos primates y explicar, por ejemplo, sus respuestas ante situaciones de inequidad sosteniendo que han formado un contenido evaluativo como *esto es inadecuado*, estamos realmente atribuyéndoles un tipo de cognición más compleja que aquella a la que debe apelar el escéptico que explica estas repuestas en términos de contenidos sofisticados, como el deseo de *recibir lo mismo que otros por la misma tarea*, o el de *recibir una recompensa equitativa*.<sup>6</sup>

5. Un problema ulterior a discutir es si hemos de otorgar a estos componentes de las evaluaciones normativas primitivas un carácter conceptual o no conceptual. De modo general, entre quienes suelen defender que el pensamiento de los animales es de naturaleza no conceptual cabe mencionar a (Evans, 1982; Peacocke, 2001; Bermúdez, 2003). Entre quienes, en cambio, se inclinan por atribuir algunas habilidades conceptuales básicas a (al menos algunas) especies de animales se encuentran (Allen y Hauser, 1991; Allen, 1999; Glock, 2000; Newen y Bartels, 2007; Camp, 2009; Carruthers, 2009; Duhau, 2011; Danón, 2013; Nelson, 2020; Monsó, 2022). Un trabajo ulterior radicaría en evaluar cuál de estas alternativas es la más adecuada en el caso de las evaluaciones normativas que aquí nos ocupan.

6. El atractivo de una explicación que apele a un tipo de normatividad primitiva como la aquí presentada aumenta, creo, en el caso del chimpancé que rechaza una recompensa más valiosa cuando su compañero recibe una recompensa menor. La razón es la siguiente: el escéptico no puede explicar esta conducta apelando a los mismos deseos egoístas que atribuía al primate que reaccionaba negativamente cuando él recibía una recompensa menor. Parece, en cambio, que tendrá que atribuirle un deseo con un contenido general como el de *que quien haya realizado un mismo trabajo reciba una recompensa equitativa*, o algo

El segundo ejemplo empírico que me interesa retomar es el de los comportamientos de control (*policing*) en primates no humanos. También en este caso, los escépticos han elaborado algunas explicaciones no normativas que apuntan a lidiar con estos comportamientos. Schmidt y Rakoczy (2019) sugieren, por ejemplo, que los primates llevan a cabo estas intervenciones porque sienten disgusto ante los comportamientos agresivos de otros y no porque los evalúen como incorrectos a la luz de algún tipo de norma. Ahora bien, esta explicación no da cuenta de un hecho evidente: no todos los primates intervienen imparcialmente ante situaciones de agresión entre terceros. Solo los miembros dominantes del grupo lo hacen. Lo cual nos lleva a preguntarnos: ¿es que sólo los miembros dominantes sienten el disgusto al que hacen referencia Schmidt y Rakoczy? Y, de no ser así, ¿de qué modo explicamos que sólo ellos intervengan para detener estos conflictos?

Schlingloff y Moore (2017) ofrecen, por su parte, otra explicación deflacionada de estas intervenciones imparciales que sí permite responder a esta pregunta. Centrándose en el caso de los chimpancés, estos autores argumentan que sería “prematureo” interpretar las intervenciones imparciales por parte de los machos dominantes como evidencia de que están reforzando en otros miembros del grupo ciertas normas, como la de *no involucrarse en peleas con otros*. La razón, nos dicen, es que los chimpancés pueden sencillamente preferir *no vivir en grupos inestables y plagados de conflictos*. Luego, si un chimpancé que debido a su rango jerárquico está en condiciones de realizar una intervención imparcial exitosa se encuentra ante una situación de conflicto, se verá motivado a intervenir de modo pacificador. Esto permite explicar los comportamientos de control en términos de las preferencias de los chimpancés y del lugar en la jerarquía que ocupan, sin involucrar ningún tipo de aprehensión de normas.

Ahora bien, aunque cuente con esta virtud explicativa, parece claro que la explicación de Schlingloff y Moore debe apelar a algunas capacidades cognitivas y representacionales relativamente complejas. Por una parte, ha de atribuir a los primates contenidos sofisticados como el deseo de *vivir en grupos pacíficos o carentes de conflictos*. Por otra, debe adscribirles alguna capacidad para creer, o registrar de algún modo, que él (y no otros) es *quien está en condiciones de intervenir exitosamente*

semejante. Quien apela a la noción de normatividad primitiva puede, en cambio, atribuir al chimpancé un mismo tipo de contenido perceptual y evaluativo como *esto no es adecuado*, que guía su conducta de rechazo tanto cuando es él, como cuando es su compañero, quien recibe una recompensa menor.



para calmar al grupo. Todo lo cual supone que estos primates poseen el concepto de *grupo*, son capaces de representarse a un grupo como *pacífico* o *conflictivo*, etc.

Como en el caso de las respuestas negativas ante la inequidad, el escéptico aún puede argüir que estos requisitos son menos demandantes que los que el intelectualista impone para considerar que la conducta de estos animales está guiada por una sensibilidad normativa. Después de todo, la explicación de Schlingloff y Moore prescinde de la cognición de segundo orden, del concepto de norma, etc.

Sin embargo, nuevamente, nos encontramos con que no es tan sencillo apelar a estas consideraciones cuando comparamos las explicaciones de Schlingloff y Moore con una explicación que apele a la noción de normatividad primitiva. Pues, en este segundo caso, podemos sencillamente pensar que el macho dominante que observa un conflicto se forma una representación como *esto no es adecuado* y, en virtud de que reconoce su propio lugar de dominancia en la jerarquía, interviene para detenerlo. Esto supone, sin duda, que este primate ha de ser capaz de percibir situaciones particulares de conflicto y evaluarlas como no adecuadas. Más aún, ha de contar con alguna disposición a detener estas situaciones de un modo específico que evalúa como el adecuado (esto es, sin tomar partido por una parte ni por otra).<sup>7</sup> Pero no es preciso que posea conceptos como el de *conflicto*, o el de *grupos pacíficos*, que figuran en los contenidos desiderativos que postulan Schlingloff y Moore. Con lo cual, nuevamente, se nos presenta la pregunta: ¿es realmente más parsimoniosa una explicación que invoque estados intencionales como el deseo, junto a un conjunto de conceptos complejos, que otra que postule capacidades para percibir y evaluar algunas situaciones particulares en tanto apropiadas o no apropiadas?

Responder a estos interrogantes requiere discusiones en detalle con respecto a las capacidades para percibir, evaluar y conceptualizar el entorno que poseen distintas especies de primates, y a la mayor o menor complejidad cognitiva que estas supongan. Lo dicho hasta aquí no basta,

7. Nótese que si apelamos a la noción de normatividad primitiva podemos explicar tanto el hecho de que los primates intervengan (porque les parece que eso que observan es inadecuado), como el modo imparcial en que lo hacen (porque les parece que esto que hacen es lo adecuado). Una apelación a estados mentales intencionales no normativos, en cambio, puede explicar la intervención del primate dominante postulando que *le disgusta* la situación de conflicto que observa. Sin embargo, no parece sencillo explicar en estos términos por qué actúa imparcialmente, ya que, aparentemente, para responder de este modo el chimpancé debe ir en contra de sus inclinaciones naturales a ayudar a sus parientes y aliados (incluyendo aquí sus preferencias o deseos).

por lo tanto, para establecer que las explicaciones en términos de normatividad primitiva son superiores a sus rivales a la hora de dar cuenta de comportamientos sociales complejos de los primates, como las intervenciones pacificadoras imparciales o las respuestas a la inequidad. Sin embargo, cabe insistir en que las explicaciones que podemos producir al apelar a la noción de normatividad primitiva solo postulan capacidades perceptivas y evaluativas de primer orden, y contenidos concretos acerca de sucesos o estados de cosas que tienen lugar en el entorno inmediato de las criaturas. Esto las convierte, *prima facie*, en rivales explicativas prometedoras, que no pueden descartarse rápidamente apelando al Canon de Morgan.

La discusión anterior debería conducirnos, pienso, a extraer dos conclusiones. La primera, es que no debemos minimizar la relevancia que tiene, a la hora abordar el problema de la normatividad animal, detenernos a examinar de qué modo, entre varios posibles, estamos entendiendo la sensibilidad a las normas. Pues, como vimos, esto incidirá en cuán persuasivas resulten ciertas explicaciones, de carácter no normativo, que se puedan ofrecer de estos mismos comportamientos. La segunda es que, si nos interesa explorar la posibilidad de atribuir a los primates no-humanos una sensibilidad primitiva a las normas, tendremos que debatir en detalle los requisitos cognitivos que esta involucra y refinar nuestra búsqueda de aquella evidencia empírica que nos permita dirimir si algunos de estos animales pueden satisfacerlos. En tanto estas tareas sigan pendientes, cualquier conjetura con respecto a si hay primates que son sensibles a las normas es, como nos advertía el propio de Waal, tentativa. Lo cual no nos impide empezar a trazar distinciones conceptuales que nos ayuden a entender qué está en juego cuando nos preguntamos si hay algún tipo de cognición normativa que compartamos con nuestros parientes evolutivos más cercanos: los primates no humanos.

## REFERENCIAS

- Allen, Colin (1999). Animal Concepts Revisited: The use of Self-Monitoring as an Empirical Approach. *Erkenntnis*, 51(1), 537-544. <https://doi.org/10.1023/A:1005545425672>
- Allen, Colin y Hauser Marc (1991). Concept Attribution in Non-Human Animals: Theoretical and Methodological Problems in Ascribing Complex Mental Processes. *Philosophy of Science*, 58(2), 221-240. <https://doi.org/10.1086/289613>

- Andrews, Kristin (2020). Nāive Normativity: The Social Foundation of Moral Cognition. *Journal of the American Philosophical Association*, 6(1), 35-56. <https://doi.org/10.1017/apa.2019.30>
- Audi, Robert (2018). Moral Perception Defended. En Anna Bergqvist y Robert Cowan (Eds.), *Evaluative Perception* (pp. 58-79). Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198786054.003.0004>
- Aureli, Filippo y de Waal, Frans (Eds.) (2000). *Natural Conflict Resolution*. Berkeley, CA: University of California Press. <https://doi.org/10.1525/9780520924932-003>
- Barrett, Clark (2015). *The Shape of Thought. How Mental Adaptations Evolved*. Nueva York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199348305.001.0001>
- Bergqvist, Anna y Cowan Robert (Eds.) (2018). *Evaluative Perception*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198786054.001.0001>
- Bermúdez, Jose Luis (2003). *Thinking Without Words*. Nueva York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195159691.001.0001>
- Boehm, Christopher (1994). Pacifying Interventions at Arnhem Zoo and Gombe. En Richard Wrangham, William McGrew, Frans de Waal y Paul Heltne (Eds.), *Chimpanzee Cultures* (pp. 211-226). Cambridge MA: Harvard University Press.
- Brosnan, Sarah y de Waal, Frans (2003). Monkeys Reject Unequal Pay. *Nature*, 425(6955), 297-299. <https://doi.org/10.1038/nature01963>
- Brosnan, Sarah, Schiff, Hillary, y de Waal, Frans (2005). Tolerance for Inequity may Increase with Social Closeness in Chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, 272(1560), 253-258. <https://doi.org/10.1098/rspb.2004.2947>
- Brosnan, Sarah, Talbot, Catherine, Ahlgren, Megan, Lambeth, Susan y Schapiro, Steven (2010). Mechanisms Underlying Responses to Inequitable Outcomes in Chimpanzees, *Pan troglodytes*. *Animal Behaviour*, 79(6), 1229-1237. <https://doi.org/10.1016/j.anbehav.2010.02.019>
- Burge, Tyler (2010). *Origins of Objectivity*. Nueva York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199581405.001.0001>
- Camp, Elisabeth (2009). Putting Thoughts to Work: Concepts, Stimulus-Independence and the Generality Constraint. *Philosophy and Phenomenological Research*, 78(2), 275-31. <https://doi.org/10.1111/j.1933-1592.2009.00245.x>
- Carruthers, Peter (2009). Invertebrates Concepts Confront the Generality Constraint (and Win). En R. Lurz (ed.), *The Philosophy of Animal Minds*

- (pp. 89-107). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511819001.006>
- Cords, Marina y Aureli, Filippo (2000). Reconciliation and Relationship Qualities. En Filippo Aureli y Frans de Waal (eds.), *Natural Conflict Resolution* (pp. 177-198). Berkeley, CA: University of California Press, Berkeley. <https://doi.org/10.1525/9780520924932-014>
- Danón, Laura (2013). Conceptos de sustancias y conceptos de propiedades en animales no humanos. *Crítica*, 45(133), 27-54. <https://doi.org/10.22201/iifs.18704905e.2013.712>
- Danón, Laura (2019). Animal Normativity. *Phenomenology and Mind*, 17, 176-187.
- de Waal, Frans (1986). The Integration of Dominance and Social Bonding in Primates. *The Quarterly Review of Biology*, 61(4), 459-479. <https://doi.org/10.1086/415144>
- de Waal, Frans (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge/Londres: Harvard University Press. <https://doi.org/10.4159/9780674033177>
- de Waal, Frans (2000). Primates: A Natural Heritage of Conflict Resolution. *Science*, 289(5479), 586-590. <https://doi.org/10.1126/science.289.5479.586>
- de Waal, Frans (2014). Natural Normativity: The 'is' and 'ought' of Animal Behavior. *Behaviour*, 152(2,3), 185-204. [https://doi.org/10.1163/9789004263888\\_006](https://doi.org/10.1163/9789004263888_006)
- de Waal, Frans y Roosmalen, Angeline (1979). Reconciliation and Consolation Among Chimpanzees. *Behavioral Ecology and Sociobiology*, 5(1), 55-66. <https://doi.org/10.1007/BF00302695>
- Dubreuil, Diane, Gentile, María Silvia, y Visalberghi, Elisabeth (2006). Are Capuchin Monkeys (*Cebus apella*) Inequity Averse? *Proceedings of the Royal Society B*, 273(1591), 1223-1228. <https://doi.org/10.1098/rspb.2005.3433>.
- Duhau, Laura (2011). Do Bees Really Have Concepts? A Discussion of Carruthers' Criteria of Conceptuality. *Teorema*, 29(2), 125-134.
- Evans, Gareth (1982). *The Varieties of Reference*. New York: Oxford University Press.
- Fitzpatrick, Simon (2017). Against Morgan's Canon. En Kristin Andrews y Jacob Beck (Eds.), *Routledge Handbook of Philosophy of Animal Minds* (pp. 437-447). London/New York: Routledge. <https://doi.org/10.4324/9781315742250-43>
- Fitzpatrick, Simon (2020). Chimpanzee Normativity: Evidence and Objections. *Biology and Philosophy*, 35(4), 1-18. <https://doi.org/10.1007/s10539-020-09763-1>

- Flack, Jessica, de Waal Frans y Krakauer David (2005). Social Structure, Robustness, and Policing Cost in a Cognitively Sophisticated Species. *The American Naturalist*, 165(5), E126-E139. <https://doi.org/10.1086/429277>
- Flack, Jessica, Girvan, Michelle, de Waal, Frans, y Krakauer, David (2006). Policing Stabilizes Construction of Social Niches in Primates. *Nature*, 439(7075), 426-9. <https://doi.org/10.1038/nature04326>
- Flack, Jessica, Jeannotte Lisa y de Waal Frans (2004). Play Signaling and the Perception of Social Rules by Juvenile Chimpanzees (*Pan Troglodytes*). *Journal of Comparative Psychology*, 118(2), 149-159. <https://doi.org/10.1037/0735-7036.118.2.149>
- Ginsborg, Hannah (2011). Primitive Normativity and Skepticism about Rules. *Journal of Philosophy*, 108(5), 227-254. <https://doi.org/10.5840/jphil2011108518>
- Ginsborg, Hannah (2018). Normativity and Concepts. En Daniel Stars (Ed.), *The Oxford Handbooks of Reasons and Normativity* (pp. 989-1014). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhdb/9780199657889.013.43>
- Glock, Hans-Johann (2000). Animals, Thoughts and Concepts. *Synthese*, 123(1), 35-64. <https://doi.org/10.1023/A:1005295521736>
- Glock, Hans-Johann (2019). Agency, Intelligence and Reasons in Animals. *Philosophy*, 94(4), 645-671. <https://doi.org/10.1017/S0031819119000275>
- Hopper, Lydia, Lambeth, Susan, Bernacky, Bruce, Schapiro Steven y Brosnan, Sarah (2013). The Ontogeny of Social Comparisons in Rhesus Macaques (*Macaca mulatta*). *Journal of Primatology*, 2(109), 2. <https://doi.org/10.4172/2167-6801.1000109>
- Kappeler, Peter, Fichtel, Claudia, y van Schaik, Carel (2019). There Ought to be Roots: Evolutionary Precursors of Social Norms and Conventions in Non-human Primates. En Neil Roughley y Kurt Bayertz (Eds.), *The Normative Animal? On the Anthropological Significance of Social, Moral and Linguistic Norms* (pp. 65-82). Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190846466.003.0003>
- Kitcher, Philip. (2006). Morality and Evolution. How to Get from Here to There. En Stephen Macedo y Josiah Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 120-139). Princeton NJ: Princeton University Press. <https://doi.org/10.1515/9781400830336-009>
- Korsgaard, Christine (2006). Morality and the Distinctiveness of Human Action. En Stephen Macedo y Josiah Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 98-119). Princeton, NJ:

- Princeton University Press. <https://doi.org/10.1515/9781400830336-008>
- Korsgaard, Christine (2010). Reflections on the Evolution of Morality. *The Amherst Lecture in Philosophy*, 5, 1-29. <http://www.amherstlecture.org/korsgaard2010>
- Lorini, Giuseppe (2018). Animal Norms: An Investigation of Normativity in the Non-human Social World. *Law, Culture and Humanities*, 18(3), <https://doi.org/10.1177/174387211880000>
- Massen, Jorg, Van Den Berg, Lisette, Spruijt, Berry, y Sterck, Elisabeth (2012). Inequity Aversion in Relation to Effort and Relationship Quality in Long-tailed Macaques (*Macaca fascicularis*). *American Journal of Primatology*, 74(2), 145-156. <https://doi.org/10.1002/ajp.21014>
- McGeer, Victoria y Pettit, Philip (2002). The Self-Regulating Mind. *Language and Communication*, 22(3), 281-299. [https://doi.org/10.1016/S0271-5309\(02\)00008-3](https://doi.org/10.1016/S0271-5309(02)00008-3)
- Monsó, Susana (2022). How to Tell If Animals Can Understand Death. *Erkenntnis*, 87(1), 117-136. <https://doi.org/10.1007/s10670-019-00187-2>
- Morgan, Lloyd (1894). *An Introduction to Comparative Psychology*. Londres: The Walter Scott Publishing Co. Ltd. <https://doi.org/10.1037/11344-000>
- Neiworth, Julie, Johnson, Elizabeth, Whillock, Katherine, Greenberg, Julia, y Brown, Vanessa (2009). Is a Sense of Inequity an Ancestral Primate Trait? Testing Social Inequity in Cotton Top Tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*, 123(1), 10-17. <https://doi.org/10.4172/2167-6801.1000109>
- Nelson, Erik (2020). What Frege asked Alex the Parrot: Inferentialism, Number Concepts, and Animal Cognition. *Philosophical Psychology*, 33(2), 206-227. <https://doi.org/10.1080/09515089.2019.1688777>
- Newen, Albert y Bartels, Andreas (2007). Animal Minds and the Possession of Concepts. *Philosophical Psychology*, 20(3), 283-308. <https://doi.org/10.1080/09515080701358096>
- Nogueira de Carvalho, Felipe (2016). *Demonstrative Thought: A Pragmatic View*. Berlín/Boston: De Gruyter. <https://doi.org/10.1515/9783110465808>
- Oberliessen, Lina, Hernandez-Lallement, Julien, Schäble, Sandra, van Wingerden, Marjin, Seinstra, Maayken, y Kalenscher, Tobias (2016). Inequity Aversion in Rats, *Rattus norvegicus*. *Animal Behaviour*, 115, 157-166. <https://doi.org/10.1016/j.anbehav.2016.03.007>
- Okrent, Mark (2018). *Nature and Normativity: Biology, Teleology and Meaning*. Nueva York: Routledge. <https://doi.org/10.4324/9781315276700>
- Palagi, Elisabetta (2007). Play at Work: Revisiting Data Focusing on Chimpanzees (*Pan troglodytes*). *Journal of Anthropological Sciences*, 85, 63-81.

- Peacocke, Christopher (2001). Phenomenology and Non-conceptual Content. *Philosophy and Phenomenological Research*, 62(3), 609-615. <https://doi.org/10.1111/j.1933-1592.2001.tb00077.x>
- Price, Sara y Brosnan, Sarah (2012). To Each According to his Need? Variability in the Responses to Inequity in Non-human Primates. *Social Justice Research*, 25(2), 140-169. <https://doi.org/10.1007/s11211-012-0153-z>
- Range, Friederike, Horn, Lisa, Viranyi, Zsófia, y Huber, Ludwig (2009). The Absence of Reward Induces Inequity Aversion in Dogs. *Proceedings of the National Academy of Sciences*, 106(1), 340-345. <https://doi.org/10.1073/pnas.0810957105>
- Ren, Renmei, Yan Kanghui, Su, Yanjie, Qi, Hanjun, Liang, Bin, Wunyong, Bao y de Waal, Frans (1991). The Reconciliation Behavior of Golden Monkeys (*Rhinopithecus roxellanae roxellanae*) in Small Breeding Groups. *Primates*, 32, 321-327. <https://doi.org/10.1007/BF02382673>
- Roma, Peter, Silberberg, Alan, Ruggiero Angela, y Suomi, Stephen. Capuchin monkeys, Inequity Aversion, and the Frustration Effect (2006). *Journal of Comparative Psychology*, 120(1), 67-73. <https://doi.org/10.1037/0735-7036.120.1.67>. PMID: 16551166.
- Rudolf von Rohr, Claudia, Burkart, Judith y van Schaik Carel (2010). Evolutionary Precursors of Social Norms in Chimpanzees: A New Approach. *Biology and Philosophy*, 26(1), 1-30. <https://doi.org/10.1007/s10539-010-9240-4>
- Rudolf von Rohr, Claudia, Koski, Sonja, Burkart, Judith, Caws, Clare, Fraser, Orlaith, Ziltener, Angela y van Schaik, Carel (2012). Impartial Third-Party Interventions in Captive Chimpanzees: A Reflection of Community Concern. *PLOS ONE*, 7(3). <https://doi.org/10.1371/journal.pone.0032494>
- Ryle, Gilbert (1949/2002). *The Concept of Mind*. Chicago: University of Chicago Press.
- Schlingloff, Laura y Moore Richard. (2017). Do Chimpanzees Conform to Cultural Norms? En Kristin Andrews y Jacob Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* (pp. 381-389). Nueva York: Routledge. <https://doi.org/10.4324/9781315742250-37>
- Schmidt, Marco y Rakoczy, Hannes (2019). On the Uniqueness of Human Normative Attitudes. En Kurt Bayertz y Neil Roughley (Eds.), *The Normative Animal? On the Anthropological Significance of Social, Moral and Linguistic Norms* (pp. 121-135). Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190846466.003.0006>
- Sellars, Wilfrid (1954). Some Reflections on Language Games. *Philosophy of Science*, 21(3), 204-228. <https://doi.org/10.1086/287344>

- Silberberg, Alan, Crescimbene, Lara, Addesi, Elsa, James, Anderson y Elisabetta Visalberghi (2009). Does Inequity Aversion Depend on a Frustration Effect? A Test with Capuchin Monkeys (*Cebus apella*). *Animal Cognition*, 25(3), 707-708. <https://doi.org/10.1007/s10071-021-01578-1>
- Suddendorf, Thomas (1999). The Rise of the Meta-mind. En Michael Corballis y Stephen Lea (Eds.), *The Descent of Mind: Psychological Perspectives on Hominid Evolution* (pp. 218-260). Oxford: Oxford University Press.
- Tajima, Tomoyuki y Kurotori, Hidetoshi (2010). Nonaggressive Interventions by Third Parties in Conflicts among Captive Bornean Orangutans (*Pongo pygmaeus*). *Primates*, 51(2), 179-182. <https://doi.org/10.1007/s10329-009-0180-z>.
- Vervaecke, Hilde, de Vries, Han, y van Elsacker, Linda (2000). Function and Distribution of Coalitions in Captive Bonobos (*Pan paniscus*). *Primates*, 41(3), 249-265. <https://doi.org/10.1007/BF02557595>
- Vincent, Sara, Ring, Rebeca y Andrews, Kristin (2018). Normative Practices of Other Animals. En Aaron Zimmerman, Karen Jones y Mark Timmons (Eds.), *The Routledge Handbook of Moral Epistemology* (pp.57-83). Nueva York: Routledge. <https://doi.org/10.4324/9781315719696-4>
- Wascher, Claudia y Bugnyar, Thomas (2013). Behavioral Responses to Inequity in Reward Distribution and Working Effort in Crows and Ravens. *PLOS ONE*, 8(2). <https://doi.org/10.1371/journal.pone.0056885>
- Watts, David, Colmenares, Fernando y Arnold, Kate (2000). Redirection, Consolation and Male Policing: How Targets of Aggression Interact with Bystanders. En Aureli, Filippo y Frans de Waal (Eds.), *Natural Conflict Resolution* (pp.281-301). Berkeley, CA: University of California Press. <https://doi.org/10.1525/9780520924932-020>
- Westra, Evan y Andrews, Kristin (2022). A Pluralistic Framework for the Psychology of Norms. *Biology and Philosophy*, 37(5),1-30. <https://doi.org/10.1007/s10539-022-09871-0>
- Wolkenten, Megan, Brosnan, Sarah y de Waal, Frans (2007). Inequity Responses of Monkeys Modified by Effort. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18854-18859. <https://doi.org/10.1073/pnas.0707182104>



eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.31453>

## JUSTICE AND RELATED MATTERS IN THE LEGACY OF FRANS DE WAAL

### *La justicia y otras cuestiones afines en el legado de Frans de Waal*

Vicente CLARAMONTE SANZ   
*Departamento de Filosofía, Unidad Docente de Lógica y Filosofía de la  
Ciencia, Universitat de València, España*  
[vicente.claramonte@uv.es](mailto:vicente.claramonte@uv.es)

Rodolfo GUARINOS RICO  
*Departamento de Filosofía, Unidad Docente de Lógica y Filosofía de la  
Ciencia, Universitat de València, España*  
[rodolguarinos@gmail.com](mailto:rodolguarinos@gmail.com)

Recibido: 24/05/2023    Revisado: 22/07/2023    Aceptado: 23/07/2023

**ABSTRACT:** Initially, we critically examine the current state of knowledge in the field of Philosophy of Biology pertaining to empathy as a subject of scientific investigation. Subsequently, we delve into the constraints associated with empathy in some primates, including humans, recognizing it as a socially situated and evolved attribute. Additionally, we explore its potential as a political asset among humans. In light of these findings, we reevaluate the dichotomy between perspectives that emphasize altruism and egoism as fundamental principles in the biological and ontological senses. Lastly, we put forth the scientifically defensible

idea that significant connections exist between empathy and the concept of justice.

*Keywords:* scientific iusnaturalism, altruism, egoism, evolutionary epistemology, Philosophy of Biology.

**RESUMEN:** El artículo comienza analizando el estado actual del debate en Filosofía de la Biología sobre el establecimiento de la empatía como objeto del conocimiento científico. A partir de esta premisa, reflexiona acerca de los límites de la empatía en la configuración de la conducta humana planteando ciertos argumentos conducentes a una revisión del equilibrio de fuerzas en el debate entre el altruismo biológico y el egoísmo ontológico. Finalmente, propone la posibilidad de afirmar la existencia de nexos científicamente demostrables entre la aptitud para la empatía y la idea de justicia.

*Palabras clave:* iusnaturalismo científico, altruismo, egoísmo, epistemología evolucionista, Filosofía de la Biología.

## 1. INTRODUCTION

Frans de Waal's work has been causing a significant impact in the field of Philosophy of Biology for decades. This impact is not only due to the valuable intrinsic content of his systematic empirical observations and theoretical contributions—such as politics in chimpanzees, empathy as an object of scientific knowledge, the natural genesis of Ethics, and aversion to inequality—but also because it triggers questions and critical issues that will likely encourage a reexamination of previously widely accepted or presupposed notions. It may also lead to a shift in the analytical perspective to be applied in the lines of research opened by his work for future investigations. Attempting to condense, describe, or even evaluate over forty years of rigorous scientific labor in a few pages would be unwise, if at all possible. Fortunately, the ambitions of this text are much more modest. Nonetheless, we would like to propose, through this text, some reflections that can contribute to the debate on the consequences that de Waal's intellectual legacy is promoting in terms of changing perspectives or questioning certain dogmas or axioms presupposed in the preceding philosophical-biological paradigm, as we believe they should be reevaluated in light of his work. Let us now explore how we can suggest some reflections in this regard.

## 2. THE ESTABLISHMENT OF EMPATHY AS AN OBJECT OF SCIENTIFIC KNOWLEDGE

If Sigmund Freud underlined the importance of the concept of the unconscious within the field of Psychology, it could be said that De Waal has established the visibility of empathy as an object of scientific knowledge within Biology. In doing so, he has strengthened contemporary evolutionary epistemology. Thanks to his research, emotions have ceased to be opaque, invisible, or negligible psychic and behavioural phenomena for science and have attained the epistemological status of a subject of study as intrinsic to scientific-biological knowledge as, and connected with, species' evolutionary trajectories, hereditary transmission through genetics, epigenetics, etc., the particular and differentiable action of evolutionary forces, or physiology. After all, empathy is comprehended as the cognitive ability to identify oneself with others and share their emotions or feelings. It is the capacity of the self to be affected by the other-selves, involving the internalization of others' psychological states to the extent of experiencing them—or nearly so—as if they were one's own. Ultimately, it represents a remarkable and precise degree of emotional resonance.

What would be the meaning of emotions in Biology, paraphrasing Theodosius Dobzhansky (1973), in light of evolution? What is their teleology if we adopt a philosophical-biological perspective based on the notion of final cause as proposed by Aristotle? Initially, emotions seem a sort of interface between three distinct entropic systems, although two of them are strongly intertwined: mind, body, and the environment (De Waal, 2019, 112). Instincts also represent an organism's response to the interactions between mind, body, and environment. However, in the vast majority of animal species, instincts entail an automatic and unambiguous reflexive response. When a stimulus is received in this triple interface, an immediate, predetermined, and closed response is activated. This circumstance was previously considered by authors like Descartes (1991) as a premise to deem all non-human animals as automata or machines, at most, organic matter subject to the rigid and inexorable laws of Physics and Chemistry but lacking consciousness of their own individual identity.

Although emotions, like instincts, also direct the mind's attention and predispose the body to promote actions and ultimately behaviours, emotions emerge without the automatism and immediacy characteristic of instinctive responses. They allow for a space—time—for the prior meditation of experience, reflection, and judgment, thus constituting a flexible and open psychic system much more sophisticated and richer than

instincts. Instead of the previously described unidirectional, one-to-one relationship between the stimulus from the environment and the reaction of the mind-body entity that constitutes the organism, which is characteristic of instincts, emotions enable a multivalent behavioural response. They allow for the integration of individual experience and the evaluation of the environment, thereby preparing the organism for an optimal response. From the perspective of an evolutionary epistemology, it could be argued that emotions evolved because they decisively contribute to guiding our behaviour in an environment of elevated complexity that cannot be fully comprehended cognitively. Specifically, they do so by their capacity to induce adaptive reactions based on the evaluation of various situations imposed by the environment, such as danger, the unknown, competition for limited resources in a given environment, external aggressions, mating options, and more (De Waal, 2019, 35).

Ultimately, emotions are the physiological-metabolic mechanism through which evolution ensures that, in intensely social species with a high degree of cognitive development such as those included in the primate taxon, the body implements the most beneficial responses and behaviors for the adaptive success of the organism in each environment and under each circumstance. While it is true that instinct also plays a similar role in many socially and cognitively less complex species, as mentioned above, the key advantage of emotions lies in the fact that they “do not dictate specific behaviors” (De Waal, 2019, 113). They are not automatic, one may add, and thus allow for the mediation of factors that are prototypical of *Homo sapiens*, such as experience, deliberation, prudence, reasoned judgment, and so on. It is precisely because our behavioral repertoire is so similar to that of other primates in many aspects that we, primates, experience similar emotions.

Emotions are closely tied to the evaluative cognition of the environment, constantly categorizing information to optimize our adaptability and, ultimately, ensure our survival. If emotions were solely the result of indulging in our own and others' moods, they would likely lack any evolutionary significance and cease to exist. Since Rizzolatti and Sinigaglia's (2006) groundbreaking work, we have abundant scientific evidence to assert that emotional empathy is not a supernatural essence, a kind of “ghost in the machine” harbored exclusively by our privileged species due to its unique sensitivity. Instead, it has a well-documented physiological basis responsible for its efficacy: mirror neurons. These neuro-specular mechanisms enable resonance between different selves and the simulation of others' psychological states “as if” they were our own. Furthermore, thanks to documented cases such as Phineas Gage and Elliot, as studied

by Antonio Damasio (2001), and subsequent research by neuroscientists like Marco Iacoboni (2009), it can be confidently stated that emotions not only constitute the *conditio sine qua non* of moral reasoning but also serve as a prior and crucial element for almost any type of decision-making. In fact, without emotions involved in the various available options, reflection and reasoning alone are insufficient for making choices.

In other words, even if any specimen of *Homo sapiens* retains intact the capacity for abstract reasoning, coherent analysis, strategy, and calculation, without emotions, their ability to make decisions and, consequently, to make any moral judgment declines. Brain scanning techniques and positron emission tomography have repeatedly demonstrated that the resolution of moral conflicts activates evolutionarily ancient brain areas. Therefore, moral decision-making cannot depend solely on the extended neocortex, our most recent acquisition in brain structures responsible for abstract reasoning. Instead, it is firmly rooted in “millions of years of social evolution” (De Waal, 2010, 43). For the same reasons, without emotions, human beings would lack any cognitive evaluation of the environment that allows us to optimize our adaptability and increase our chances of survival.

Given that emotions have played a decisive role in our evolutionary history, both because they are essential for cognitive evaluation of the environment and because they constitute the indispensable link in the development of the intense sociability characteristic of *Homo sapiens*, why should emotions not be subject to scientific inquiry in natural sciences, just like other human cognitive abilities such as memory, perception, attention, or imagination?

### 3. LIMITS OF EMPATHY IN HOMO SAPIENS: A CONTEXTUALIZED ABILITY

To admit empathy as an object of study in natural sciences, and by extension in Philosophical Anthropology, implies in some way drawing consequences from its application to the analysis of our own behaviour and our philosophical-political conception of the globalized social world in which we live. Empathy affects the very core of a longstanding debate that emerged shortly after the advent of Darwinism in Biology, the dialectic between selfishness and altruism as an explanatory principle for individual and collective behaviour, and ultimately for the worldview that shapes the human political community, as it is closely linked to the notion of altruism and, more specifically, with the very possibility of experiencing and sharing it. Once altruistic conducts are admitted as real phenomena in natural sciences, the critical question in evolutionary biology would be

how to explain it within a dynamic marked by the struggle for survival, especially if we assume individuals as entirely selfish entities compelled by competition for resources and governed by the relentless tribunal of natural selection. This is because, if we assume an ontology in which selfishness is fundamental for reproductive benefit, altruistic behaviours would be displaying an antiadaptive behaviour, evolutionarily irrational and ineffective, as they cease to prioritize the self and the satisfaction of one's own needs as the absolute priority in any circumstance.

However, doctrine does not agree when it comes to precisely establishing what constitutes altruistic behaviour, beyond generic notions. In principle, one could speak of altruistic behaviour of an organism when, on its own initiative, without receiving any external coercion or expecting compensatory reciprocity, it shares its own resources to satisfy the needs of others. However, this approach would be clearly insufficient when considering altruism as a complex phenomenon involving many internal and external factors, as well as a no less intricate evolution. In fact, various theories have attempted to explain it to date, but all have been deficient in defining altruism, lacking clarity and precision, thereby hindering the interpretation of relevant empirical data. Some portray altruism as merely apparent, driven by fundamental selfishness, «Scratch an 'altruist' and watch a 'hypocrite' bleed» (Ghiselin, 1974, 247), while others attempt to present it as genuinely selfless. A final group of options seeks to explain altruism through processes of self-identification with the other, excluding acts of supererogatory self-sacrifice and self-giving due to situating the debate outside the egoism-altruism dichotomy.

Certainly, solutions in which human motivations are inherently selfish have long been dominant. Against their inability to offer a sufficient explanation of the phenomena studied, new altruistic explanations are designed against this long-held monopoly. However, altruistic theories have problems like lack of precision leading to misinterpretation and inconclusiveness of the empirical findings, according to Feigin *et al.* (2014, 6). Given this, perhaps it is time to propose interpreting this circumstance as a symptom that “the assumption of universal egoism must be replaced by a more complex assumption allowing room for both egoism and altruism” (Feigin *et al.*, 2014, 6). And this should be done while avoiding the crude abuse of Natural History as the ultimate deciding judge to argue for or against either based on empirical observations, as nature is so rich and complex that countless observations could be cited without advancing the controversy in any direction: “The book of nature is like the Bible: everyone reads into it what they want to read, tolerance or intolerance, altruism or greed” (De Waal, 2015, 61).

In any case, it seems clear that behaviors considered altruistic are highly conditioned by the circumstances in which they occur. Moreover, it is difficult to avoid the suspicion that altruistic dispositions coevolved with other behaviors, such as the rejection of outsiders, favoritism towards one's own group, and certain types of cooperation with specific members of the same group but not others. To address altruism in its complexity, we will primarily focus on the phenomenon of parochial altruism, defined as favoritism towards the internal group at the expense of the external group (Choi and Bowles, 2007). Although this phenomenon of parochial altruism is often considered an anthropological universal of the sapiens (Greene, 2013), its frequency actually varies among communities and throughout individuals' lives, just as the empirical evidence presented in the specialized literature varies depending on the methodological approach used to study it (Pisor *et al.*, 2020). This points, on one hand, to a flexible phenomenon that is context-dependent and highly complex, and on the other hand, to the inconvenience of assuming empathy culminating in altruism as the decisive argument to definitively settle the manichean dilemma of selfishness versus altruism in favor of the latter.

In principle, the specialized literature on the aforementioned parochial altruism commonly conceives it as a primary disposition in our species, so that tolerance towards outgroup individuals would only emerge due to the suppression of this tendency towards the, let's say, natural inclination to parochialism. However, it is not entirely clear in this literature what the necessary and sufficient causes of this suppression are. Among the different possible causes, cultural institutions can impose tolerance towards outgroup members (Fearon and Laitin, 1996; Fry, 2018), the generation of new loyalties through interaction with outgroup members would reduce favoritism towards the ingroup (Brewer and Campbell, 1976; Buchan *et al.*, 2009; Fukuyama, 2001; Hruschka and Henrich, 2013; Mau *et al.*, 2008; Singer, 1981), or the adequate satisfaction of basic needs would enable individuals to genuinely consider the well-being of members from external groups and perhaps make them more inclined to do so (Hruschka *et al.*, 2014; Silva and Mace, 2014). Furthermore, each of these options has its own internal debates that further complicate the overall discussion. We may or may not accept the ontological priority of parochialism over tolerance and empathy towards outsiders, and we may even fundamentally accept the establishment of empathy and, by extension, altruism as objects of scientific knowledge. However, we must not lose sight of the intimate connections between altruism and many other human traits, some of which are often considered antagonistic to attitudes based on our empathic abilities. Perhaps this antagonism is not absolute or, if it is,

it may have an unavoidable feedback component. In any case, parochial altruism can cast light on altruism as a situated, dynamic phenomenon.

How does parochial altruism function in different human groups and in other closely related primate species such as chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*)? In these and other species, a frequent correlation has been observed between cooperation within the group and the presence of external threats. Scientific evidence shows that in different taxa, including birds and mammals, an external threat is immediately followed by an increase in group cohesion and affiliation (Radford *et al.*, 2016). Birds (Radford, 2011), social carnivores (Thompson *et al.*, 2017; Morris-Drake *et al.*, 2019; Mosser and Packer, 2009), and primates (Beehner and Kitchen, 2007) demonstrate cooperative actions among many group members before and during external conflicts.

In humans, it has been experimentally demonstrated that preferences for one's own group over external groups in competitive contexts manifest from a very early age (Fehr *et al.*, 2008; Benozio and Diesendruck, 2015; Fehr *et al.*, 2013). In chimpanzees, hostile and violent conflicts with external groups appear to be widespread, although the intensity and cost of these conflicts vary depending on the studied population (Wilson *et al.*, 2014). For example, when neighbouring chimpanzee communities extend their territorial occupation into each other's territory, they create a zone of overlap where there is a high risk of conflict between the two communities erupting (Wrangham *et al.*, 2007; Wilson *et al.*, 2007).

This territorial aspect of chimpanzee life is manifested through indirect conflicts such as vocal exchanges at a distance (Samuni *et al.*, 2021) or direct conflicts involving visual and/or physical contact, including chasing and attacking the external group (Boesch and Boesch-Achermann, 2000; Wrangham and Glowacki, 2012). These situations of fluctuating hostility can escalate to lethal aggression (Wilson *et al.*, 2014). An example of the culmination of this spiral of intraspecific aggression, known as lethal warfare, was well-documented in Gombe National Park, Tanzania, by the renowned primatologist Jane Goodall (1986), referred to in primatology as the Gombe Chimpanzee War or the Four-Year War. In this war, a chimpanzee community split into two factions, forming two separate communities. The two factions started patrolling the border area and engaging in confrontations, which escalated into an all-out war where even the elderly and revered members of the once-united community were not spared. The horrified primatologists witnessed former comrades annihilating their former friends and even drinking their warm blood.

However, this bellicose aspect of chimpanzee social life is counterbalanced by cooperative dynamics within the internal group, such as



territorial patrols and increased group cohesion (Samuni *et al.*, 2017). Similarly, it has been observed that various species of dolphins (Delphinidae) also form male coalitions that engage in collective surveillance and territorial control, displaying openly aggressive behaviors against any external group intrusion in their habitat (Parsons *et al.*, 2003).

The cases of empathy we tested do not revolve around individual otherness between the self and another self, which also exist and have been extensively documented. For example, adult chimpanzees diving into the water to rescue a young chimpanzee and ending up drowning within minutes—chimpanzees cannot swim, and they know it. Instead, we sought to focus our attention on collective responses against other groups in which some individuals risk their physical integrity to help the group. For instance, protecting shared benefits such as suitable space for all group members or the substantial amounts of protein provided by a cooperatively hunted large prey. The significance of these territorial conflicts between groups is that “territorial expansions increase feeding opportunities, reduce within-group competition, and therefore offer reproductive benefits” (Lemoine *et al.*, 2022, 6). According to Wilson *et al.* (2001), these collective responses to intrusions are characterized by loud vocalizations in chorus and patrolling, showing a group cooperation inseparable from the rejection of the external group from other chimpanzee communities.

Following Lemoine *et al.* (2022, 6) again, these and other data “confirm that chimpanzee intergroup competition, in-group cooperation and social cohesion are intimately linked.” They also suggest that “participation to border patrols involves not only immediate but also long-term fitness benefits associated with securing and potentially expanding a territory, thereby reducing within-group feeding competition and improving group members reproductive success” (Lemoine *et al.*, 2022, 7), indicating a possible connection between intra-group cooperation and intergroup competition.

Regarding bonobos, they are generally considered less xenophobic and more tolerant, as their conflicts, as far as we know, do not escalate to lethal attacks (Wilson *et al.*, 2014), and they also share territory with neighbouring groups (Lucchesi *et al.*, 2020; Samuni *et al.*, 2021). However, although these male coalitions are less frequent in bonobos compared to chimpanzees and are rather rare, there is also intergroup competition among bonobos. It has been documented a greater cooperation among males within the ingroup when directed to attacking males from the outgroup, and a reduction in aggression among ingroup members during intergroup dynamics (Tokuyama *et al.*, 2019), similar to the patterns exhibited by chimpanzees. This, along with other data, suggests that “some

degree of out-group competition is evident in bonobos, albeit to a much lesser degree than chimpanzees, and that cross-population variation in intergroup hostility occurs in bonobos” (Lemoine *et al.*, 2022, 9).

In conclusion, and returning to the previous discussion on parochial altruism, from an evolutionary perspective, it can be asserted that the cooperation of human groups with other external groups outside the strictly familial or tribal circle would have coevolved alongside hostility towards external groups as a result of selective pressure from intergroup conflicts. Of course, there are other hypotheses that emphasize the collaborative aspect to a greater extent (Hrdy, 2007; Hill, 2002), but there is strong evidence of the relationship between cooperation, cohesion, and the presence of external threats. Thus, in the context of human migration and national identity, it has been demonstrated that altruism reinforces anti-immigration attitudes when based on the belief that immigration would entail some form of harm to members of the national community, indicating that anti-immigrant individuals are often driven by altruistic motives rather than selfish attitudes: “Under certain conditions, sincere altruistic motivations can not only promote cooperation such as solving collective action problems, but can also exacerbate conflict such as in the case of immigration” (Kustov, 2021, 33).

Hence, populist anti-immigration discourses have such an easy and yet effective argument to connect with the masses and mobilize them by appealing to atavistic territorial instincts, whose trigger sooner or later responds to the scapegoating of immigrants, the quintessential other to the polis, and by extension, someone conceptualized as a non-“us” simply because they are different. In evolutionary epistemology, nationalism comes to mean the exaltation of the motto “My tribe is the best,” and in its most extremist versions, it could even imply a disturbing “My tribe is the only one.” And perhaps this does not satisfy our anthropocentric ego, but both the evolutionary proximity and the practical identity of such shared behaviours with chimpanzees and, less frequently, with bonobos, regarding the evolutionary synergy between intragroup cooperation and intergroup belligerence, help us visualize the similarities in the ultimate case of lethal aggression: war. With war, the xenophobia and contempt professed towards the other group reach the extreme of declassifying fellow members of the same species to categorize them as a different species, as species membership draws the conceptual boundary between hunting and murder or between extermination and genocide. This dislocation in intellectual representations is also shared by humans and chimpanzees as a preliminary stage to the aggression of war because when the other group is perceived as different and inferior, as a different species—Adolf

Hitler and his “Jewish problem” constituted the epitome of infamy, but examples abound—it, on the one hand, fosters self-esteem and solidarity with one’s own group, and on the other hand, makes it much easier to annihilate the other. In any case, we also coincide in the method, murder, as well as in the outcome, genocide. Before that, we proceed with dehumanization, and they with “dechimpanzification” (De Waal, 2010, 142).

Thus, established the continuities in evolutionary biology, our analysis of the plausibility of the connection between parochialism and altruism could also be traced by resorting to biochemistry. Lemoine *et al.* (2022) explore the oxytocinergic system, a physiological system with various functions that apparently modulates the activity of different brain regions involved in social interaction and could therefore play a fundamental role in the maintenance and formation of social and cooperative bonds. Although initially it may have served as the foundation for the bond between mother and offspring, in various species, including chimpanzees, it would have been co-opted as an essential component for the formation of pair bonds and, in general, various forms of companionship or affiliations with group members not linked by kinship. Interestingly and coincidentally, the oxytocinergic system is also activated when individuals face a collective threat and in competition with external groups. It has been documented that intranasal administration of oxytocin in humans during experiments conducted with the intergroup prisoner’s dilemma promotes cooperation and trust among members of the ingroup and develops a defensive disposition towards the outgroup (De Dreu *et al.*, 2010; De Dreu, 2012; Ten Velden *et al.*, 2017). In chimpanzees, this same physiological system is activated in both sexes immediately before and during a conflict and during border patrols, and it is also involved in the detection and avoidance of external groups in multiple vertebrate species. Lemoine *et al.* (2022, 9) hypothesize that “This physiological pathway probably acts by promoting pro-social behaviour and thus increasing in-group interests while, in parallel, increasing awareness of potential threat from the out-group, thereby increasing out-group hostility.”

We can also consider the role played by testosterone in cases of parochial altruism in humans (Diekhof *et al.*, 2014; Reimers *et al.*, 2015). Essentially, during intragroup conflict or competition, it has been observed that testosterone promotes altruistic and tolerant behaviour—tolerance, for example, referring to the decision not to punish an individual for breaking rules—and, in general, prosocial behaviour towards members of the ingroup among male humans, while increasing hostility towards members of the outgroup. If testosterone plays this role in male human cognition, it would mean that parochial altruism is not primarily or at least not

exclusively produced by cultural factors, even if they modulate the intensity, frequency, etc., of this phenomenon.

From this perspective, if we assume that altruism is genuine—sincere and not the camouflage of more or less immediate selfish motivations—, the capacity to generate emotions is also genuine. Therefore, altruism could be understood as one of the behavioural outcomes of that cognitive aptitude we call empathy. And both altruism and empathy would be universal or potentially universal, in the sense that they manifest respectively as behaviour and aptitude in the overwhelming majority of human beings and, generally, in the rest of primate species. However, they are not universal when it comes to designating the beneficiary of altruistic attention. It is well known that human beings can display altruistic behaviours and acts of supererogatory self-sacrifice towards relatives, friends, acquaintances, and even strangers, and we refer to the latter case as philanthropy. However, this circumstance does not imply that we are not highly selective nor that we are always willing to engage in the highest sacrificial altruism. It is far from concluding that the egoistic and altruistic facets participate in our anthropological identity in a Manichean style. If it is true that altruism evolved alongside hostility towards external groups in the context of intergroup competition, then its flexibility may have a critical limit that is not imposed by more or less ontologically selfish ideological discourses but by the evolutionary course of our species. Thus, while it is true that ontologically selfish discourses manifest a blatant inconsistency with the most basic and frequent observations of comparative ethology, the radically opposite idea would be equally untenable, according to which biological altruism would be the key and ontological egoism and parochialism would be culturally acquired defects that affect individual personality and irreversibly poison society. If such a high mutual implication between altruism, parochial behaviour, and intergroup conflict has been demonstrated, it seems inadvisable to adopt either of these radical positions exclusively. It seems unwise to collect scientific evidence with the intention of definitively and apodictically demonstrating the reign of selfishness and the nonexistence of altruism, or vice versa, when explaining the behavioural idiosyncrasy of our species.

As a corollary of the previous considerations, appealing to empathy while neglecting evolutionary, biochemical, and other closely related aspects could create a mystified image in which the extent and manner in which empathy can act as a social and political ideal end up being dangerously overvalued. An excessive emphasis on empathy can lead to an exaggeration of values intimately connected to it, such as equality, while disregarding equally important aspects of human life, such as

spontaneous hierarchies based on freedom, merit, or political struggle observed in chimpanzees and humans. Thus, in the current political arena, clearly dominated by States and large corporations associated with them, such emphasis could also persuade some individuals that the state, as an administrative machine, can embody these egalitarian values as an extension of interpersonal empathy. However, to what extent can we consider the equality, altruism, and empathy that occur spontaneously, in a face-to-face style, among members of small groups to be the same as those found between the small political groups that control the legislative and executive branches of the State and the millions of governed people, with which politicians necessarily interact as if they were faceless taxes and votes providers?

It is true that Lemoine and his colleagues mention the hypothesis that the oxytocinergic system may have undergone a functional expansion during our evolutionary history as a modulator of large-scale human cooperation with both familiar and unfamiliar individuals by activating cooperation as a response to external threats. These functional expansions “have potentially happened several times in the evolution of vertebrates [153], like being co-opted from the regulation of parturition, lactation, and mother-infant bonds, to regulating pair bond formation and social-bonds, etc.” (Lemoine *et al.*, 2022, 9). However, even considering this fact, it is highly debatable to accept that this functional expansion could result in cooperative behaviours in the face of collective external threats among citizens without family ties and with diverse or conflicting interests at the State level, if it is even possible. It seems quite obvious that the relationship between individual citizens and the administrative machinery of the State is far from the original type of relationships in which empathy, altruism, and cooperation evolved. Furthermore, human history stubbornly teaches us about states that became totalitarian machines controlled by small oligarchies who, while articulating grave discourses based on equality, were primarily concerned with defending the interests of their own kind. These oligarchs cooperated and perhaps effectively empathized with members of their own group while treating the rest of the subjects under the power of the State as mere tools. Even today, in liberal democracies, politicians grouped in parties can be seen as a privileged group attempting to use citizens as potential voters through practical discourse and activism, as well as sources of wealth through coercive tax activities. If altruism has evolved to cooperate with individuals in dynamics of interpersonal social relationships of reasonable proximity, can we expect the State to be an appropriate tool for promoting empathy, altruism, and cooperation in dynamics of intersubjective, proximate relationships?

If we consider the Soviet Union as an extreme example of this State model, authors who were educated during its regime, such as Levina (2017), describe how the material equality imposed by the State ultimately only applied to the common people, this is, those who were not part of the nomenclature, while a privileged caste of corrupt authorities, permeating all levels of the State administration, followed different values. The most illustrative phenomenon of this situation could be observed in the daily coexistence of an official market, regulated and centralized by the Supreme Soviet, and a black market, so that the population lived the social duality characteristic of an inconsistent double standard. While in the public sphere, they upheld and even shared the discourse of egalitarian ideals and the consequent centralized planning that defined the egalitarian State, in the private sphere, they resorted to the black market as something commonplace and inevitable. The collective activity imposed by the State bureaucracy, especially if it is oligarchic, is far from resembling spontaneous, altruistic, and genuinely communal empathetic cooperation. Given the key role of proximity in altruism, it may very well be the case that the final result of this State-driven egalitarian experiments was inevitable, and not just a consequence of a bad application of political values and ideals.

#### **4. THE APTITUDE FOR EMPATHY AND THE IDEA OF JUSTICE**

Therefore, “empathy fuels our interest in others” (De Waal, 2019, 144) to the extent that it provokes emotional contagion, the assumption of others’ emotions as if they were our own, especially negative emotions, as it “has the unique property of transforming another person’s misfortune into a feeling of personal distress” (Hoffman, 1981, 133). From a biological-evolutionary perspective, it can be affirmed that empathy would never have evolved if it did not provide an adaptive advantage; it would lack value for the survival of the species if it did not grant mutual benefits to individuals whose emotions resonate, thus “contributing to the creation of a cooperative society in which individuals can rely on one another” (De Waal, 2019, 152). Empathy, thus considered, is the sensitive and cognitive capacity to generate adaptive prosocial behaviours, an emotion generator that promotes among group members the development of behaviours aimed at ensuring the stability of social coexistence.

However, considering empathy as inherently good or bad is trivial, just like intelligence or physical strength; it is a multi-purpose aptitude that “can be used for good or ill, depending on one’s intentions” (De Waal, 2019, 134). On the other hand, the product of empathy, emotions, can indeed

exhibit a moral inclination. Furthermore, Edward Westermarck (1912) distinguished between positive and negative retaliatory emotions: the positive ones constitute the empathic response of satisfaction derived from receiving a benefit, and their corresponding behavioural correlate mainly seeks reward, as in the case of gratitude or returning favours; the negative ones constitute the empathic response of resentment against actions harmful to the individual or the group, and their corresponding behavioural correlate mainly seeks punishment, as in the case of delayed revenge or the revenge system developed by chimpanzees. This final case is connected to the altruistic xenophobic attitude above mentioned, in which nationals rejected foreigners out of empathy to their co-nationals, who they presumed to be harmed economically as a result of foreigner's influence in national economy.

Both are crucial due to their prosocial nature, as they foster harmony and group stability while preventing excesses and opportunism, thus promoting morality towards normativity. Given that the ability to recognize certain societal norms of conduct and apply them to others and oneself helps in survival and prosperity, as demonstrated by evolutionary psychology, natural selection favoured groups that developed an articulated morality aimed at establishing normative patterns of social behaviour reinforced with incentive mechanisms, such as reward-punishment, more or less coercive.

However, moral emotion is still not complete morality, understood as the capacity to formulate moral judgments. Positive and negative retaliatory emotions, such as gratitude or revenge, operate within a limited orbit that is simultaneously selfish and egocentric, with the self-interest as its diameter, and thus they still fall short of the universality required by the normativity implicit in the idea of justice. Selfish in the sense of valuing one's own actions and others' actions solely based on the outcome of each action in relation to one's own self, while disregarding others' interests and the consequences of the action for the other self; and egocentric in the sense of exaggerating the exaltation of one's own self, considering it the centre of attention and general activity, and thus evaluating one's own actions and others' actions solely in terms of how one desires to be treated or how one does not wish to be treated. In contrast, in order to embrace morality compatible with moral judgment, the underlying emotion must transcend mere visceral instinct and dissociate from self-interest, from the cost-benefit calculation of the action in relation to one's own situation. Thus, the turning point between positive or negative retaliatory emotions and strictly moral emotions lies in selflessness, the only way to enable the impartiality and universality required for the abstract treatment

of good and evil implicit in moral judgments: without selflessness, there is no impartiality, and without impartiality, there is no justice. However, there is still a long way from strictly moral emotions to the notion of justice, a journey that inevitably requires cognitive aptitudes for the abstract treatment of good and evil to be accounted for.

Regarding its abstract categorization, there seems to be a notable difference between *Homo sapiens* and other primates who, like other evolutionarily distant species, show relative ease in evaluating their own and others' actions based on whether they are beneficial or detrimental to the self, but with much greater difficulty regarding others. Any notion of justice must have an origin, and the most logical place to look for it is the self; the studied individuals showed certain expectations about how they themselves should be treated, but not about how others should be treated. Once the capacity for at least an egocentric sense of fairness is established (De Waal, 2007, 77), it can later be expanded to encompass gradually intersubjective, social, or ultimately universal forms of justice. Recall, for example, how easily a stray dog flees if you simulate the intimidating gesture of reaching for a stone nearby—it quickly runs away to avoid harm to the self it has already experienced. Now, if we try to recall a behaviour of an individual of the same or a different species in which it inhibits its own action for the benefit or avoidance of harm to others, although such behaviours exist and have been documented, as in the case of Rhesus monkeys (*Macaca mulatta*, Wechkin *et al.*, 1964), we will find that it is more difficult for our memory.

While it is true that empathy and reciprocity observed in other species are not sufficient by themselves to generate the moral experience as understood among humans, it is equally true that they are essential for morality because without reciprocity between individuals in terms of emotional connection and exchange, there would be no moral human society. Thus, the key to ethological research demonstrating moral capacity in other animal species, such as elephants, dolphins, and especially non-human primates, lies in establishing the evolutionary continuity between that extrahuman or prehuman moral activity and human morality proper. Moreover, on the contrary, considering empathy as such a ubiquitous experience in human society, developing so early—new-borns are capable of accurately imitating facial expressions between 42 minutes and 72 hours of age (Gazzaniga, 2012, 173)—showing distinct neural and physiological correlates and having a well-established genetic basis, it would go against available knowledge in biology if it lacked any evolutionary continuity with other primates. Therefore, it is not unreasonable to assert the evolutionary continuity between the experience of morality as a constant



in human sociability, and that of other non-human primates and non-primate species, although the evidence for the latter seems weaker. It would be contrary to the available scientific evidence to continue denying the existence of moral empathy outside the species *Homo sapiens sapiens*.

However, so far, we have only analysed the emotional dimension of empathy, but not yet the cognitive dimension, which undoubtedly is essential to acknowledge moral normativity and ultimately the concept of justice. In other words, empathic response may not be limited to emotional contagion or resonance, such as when we yawn upon seeing someone nearby yawning or when we feel sorrowful when we see someone in distress. It can also include a cognitive or intellectual component in those cases that the literature in this context refers to as “taking the perspective of the other,” where empathic response also altruistically tunes in to the emotional state of others, but additionally triggers focused and cognitively selective or attributive behaviour that aims to meet the specific needs of the beneficiary of the altruistic behaviour.

Among the myriad of documented cases, we will limit ourselves to summarizing one well-known example provided by De Waal (2007, 55-59). Kuni, a female bonobo at the Twycross Zoo in England, caught a starling and, after being asked by a caretaker to release it, held the starling in one hand and climbed the nearest tall tree, wrapping her legs around the trunk to have both hands free to hold it. Upon reaching the highest point, she carefully unfolded her wings, one in each hand, and forcefully threw the bird towards the outside of the enclosure. This action involves empathy towards the other and compassion for suffering, but it also implies the cognitive or intellectual element we mentioned earlier because the same act would have been inappropriate with a fellow bonobo. Instead, Kuni perfectly grasped what would be suitable or good for a member of a different species.

In conclusion, it is an altruistic behaviour that adjusts to the specific needs of another. This reveals the structure of the Matryoshka dolls, used by De Waal (2003) as an analogy to present his model of the three layers of empathy. The first and most intimate level would correspond to contagion, which accounts for the automatic emotional impact and at its core, there is a perception-action mechanism that triggers immediate and even unconscious resonance between individuals. The next doll would represent cognitive empathy, through which the situation and motives behind the contagious or resonated emotions of others are evaluated. Finally, the third layer would correspond to attribution, in which the agent fully adopts the perspective of the other and engages in behaviour that satisfies their needs, as Kuni did with the starling. Perhaps we observe more complex

and sophisticated responses in *Homo sapiens*, but the evolutionary continuity of this mechanism appears beyond doubt.

How can we appreciate, in the case of the notion of justice, the first intimate layer of empathy responsible for automatic and immediate emotional contagion, the evaluative layer that assesses the situation and motives, and the final layer that corresponds to attribution, the full adoption of the other's perspective culminating in the development of a selective behaviour specifically adjusted to meet the needs of the other? We propose first tracing their manifestations, then observing their relationships and correlations, and finally arguing for some consistent explanatory inference.

Thus, the empathy responsible for automatic and immediate emotional contagion, in the context discussed here regarding the possible continuity between the normativity of human and non-human animals and its connection to the idea of justice, seems to manifest itself in the so-called "aversion to inequality" (Brosnan and De Waal, 2003; Brosnan *et al.*, 2004, 2010; methodology in De Waal, 2007, 71 *et seq.*). Anyone who has watched the famous one-minute video clip of the experiment with capuchin monkeys receiving different rewards, grapes or cucumbers, for the same task and had the opportunity to laugh at the automatic and immediate emotional reaction of the individual who becomes enraged and expresses their anger at feeling disregarded and offended for receiving a less appetizing reward knows what we are referring to. Note that the notion of "aversion to inequality," formulated in a positive sense, could also be called "affection for equality." In any case, the aforementioned studies showed how capuchin monkeys cognitively evaluate rewards in relative terms, comparing them with other available options and measuring their own effort against that of others. While this does not allow for precise labelling or classification of the underlying emotional and somatic upheaval accompanying such responses, it seems plausible to suggest that they are guided by prosocial emotions very similar to the human emotions that govern individual reactions derived from exerting efforts – "Equal pay for equal work," as the labour adage says –, gaining profits or suffering losses, and, in general, their attitude toward others based on the comparatively received treatment in exchanges and resource distribution. Also note how these studies establish as consistent scientific evidence the cognitive evaluation performed by capuchin monkeys before expressing their displeasure and fury at the unequal treatment for the same task: these are manifestations of the first two layers, corresponding respectively to automatic emotional contagion and cognitive or evaluative empathy. Thus, this affection for fair equality or aversion to unfair inequality

documented in numerous intensely social species could be described, if we may take a literary license, as a kind of “emotion of justice,” or at least an already resonant and simultaneously cognitive emotion that loudly demands notions such as fairness and justice.

In the case of attribution, understood as the full adoption of the other’s perspective until their specific need is satisfied, various behaviours have also been documented in non-human primates that manifest the latency of an intellectual representation close to the idea of justice. First, we have the observations made in cases of conflict resolution and reconciliation carried out by a mediator who lacks personal benefits in resolving the conflict. Since De Waal and Roosemalen (1979)’s work, such a substantial amount of scientifically corroborated evidence has been accumulated on this matter, both in De Waal’s work and in that of many other authors, that it is unnecessary to present an exhaustive compilation of existing bibliographic sources. The impartial mediator in the conflict perceives and evaluates the specific needs of the antagonists and the group—putting an end to reciprocal intersubjective aggression and restoring social peace—and acts accordingly without expecting any reward in return: “Here and now, equity consists of stopping the aggression between the contenders to restore group harmony: even if it’s not my problem, let’s get to work.” The “concern for the community” (De Waal, 1997, 265) or “community concern” (De Waal, 2022, 239) culminates in the implementation of prosocial behaviours that promote group harmony and, therefore, foster morality towards normativity, reinforcing the sense that in any conflict situation, there is a beneficial resolution criterion for the group that, taking into account opposed individual interests, goes beyond them.

But secondly, even more illustrative in this regard is the behaviour of control role documented in alpha males. Among chimpanzees, when a dispute arises, initially everyone tends to take the side of their relatives, allies, or friends, understood as those who frequently exchange alliances, food, grooming rituals, hugs, kisses, etc. Probably, the persistence of this unregulated dynamic would ultimately lead to total internal conflict. However, the alpha male does not tend to follow this general behaviour; instead, he interrupts fights by defending one of the parties or by acting impartially. Alpha males position themselves bristled between the contenders, engage in intimidating displays, or even physically separate them with their own arms, all until the hostility disappears. With these actions, their goal is not to favour one side in the conflict but rather to end it and restore social peace in the community. This control role and its proper performance emerge when the role of arbiter is dissociated from their own

social preferences. The evaluation of the importance and social function of the control role, as well as its cultural transmission, promote the notion of impartiality, in the sense of a selfless alignment in the heteronomous resolution of others' conflicts. Alongside the aforementioned concern for fair equality or aversion to unfair inequality, impartiality is the second prerequisite for developing the idea of justice. As mentioned above, without selflessness, impartiality cannot exist, just as without impartiality, justice cannot exist. This control role, arbiter, or judge required in conflict mediation was exemplarily performed by Luit, Frans de Waal's favourite chimpanzee. On one occasion, a dispute between two females escalated into a widespread brawl, until Luit intervened and used his strength to restore peace: "Unlike the others, he did not take sides, but anyone who continued to fight earned a slap" (De Waal, 2010, 85). In his intervention, he positioned himself above the conflicting parties and beyond his own social preferences, seeking to restore social harmony rather than favouring his relatives, friends, or allies. In other interventions, he sided with third parties, but his decisive action was not proportional to the time spent with him or the attention dedicated to grooming him; it was not conditioned by his personal interest.

But furthermore, *a fortiori*, this dissociation between the alpha's personal interests and the exercise of the arbiter role can occur not only within the same individual but also between different individuals (De Waal, 2010, 86-87). After Luit was overthrown by the coalition between Nikkie and Yeroen, during the initial phase of the subsequent four-year duumvirate, Nikkie attempted to intervene between warring parties as soon as a dispute arose. However, his intervention usually escalated the conflict rather than minimizing or resolving it, as he systematically aligned himself in favour of his friends or allies, in addition to harassing the older females by hitting them on the head. In contrast, the group always accepted the pacifying and reconciling manoeuvres of Yeroen, who ended up monopolizing the role of arbiter to such an extent that, in the second phase of his rule within the coalition, the official alpha male Nikkie didn't even gesture to intervene in case of a scuffle. This process demonstrates that the collective decides through implicit acceptance who holds the underlying authority of the arbiter, whose function does not necessarily have to coincide with that of the alpha male in the same individual. The effective performance of leadership, detached from the mere imposition of brute force, requires prior social acceptance of the arbitral function implemented by the leader. The effectiveness of their arbitration does not solely depend on force but rather on the authority transferred through implicit consensus

and majority compliance within the group. Once the role of control is dissociated from personal preferences, the political-community function of the impartial arbiter emerges, contributing to the transition from personalized and despotic power to leadership. If a collective develops the role of an impartial arbiter or judge, it seems plausible to infer that its members have at least a similar intellectual representation to our notion of justice and that there exists a generalized consensus among them regarding the goodness and necessity of individual or collective dispensation of justice for the stable and peaceful progress of social coexistence within the political community. Interestingly and coincidentally, in current law, an arbitrator is referred to as a designated judge. The community's interest in the arbiter and the skilful exercise of their role lies in the fact that arbitration offers the possibility of protecting the weak and their interests against the strong and theirs. At any given moment, anyone can find themselves in a position of weakness or be affected by the situation of a loved one. Therefore, the function of the arbiter is crucial and concerns the entire collective. By supporting the arbiter who is considered more effective by the majority, the group grants them the indispensable moral authority for their interventions to guarantee the restoration of harmony and, ultimately, the order of the community in the fairest and least violent manner possible.

In conclusion, the dissociation between the roles of the alpha male and the arbiter, along with the persistence of the functionality of heteronomous, disinterested, and impartial arbitration in the resolution of social conflicts, in our opinion, decisively contribute to substantiating an evolutionary explanation for the cultural emergence of justice administration in terms of socio-historical necessity.

## 5. CONCLUSION

The considerations presented in the previous sections, in our opinion, lead to the possibility of concluding the plausibility of various hypotheses that may contribute to enriching and clarifying the ongoing debate in this forum. In the short and medium term, the beginning of the end of the scientific paradigm that disregarded the empathic capacity of *Homo sapiens* and other animal species as objects of study seems to be occurring. The viability of this conceptual development may progressively unfold as the study of mirror neurons advances as the physiological basis of emotional empathy. This new scientific object of study, empathy, as the driving force behind social emotions and ultimately

adaptive prosocial behaviours through a more sophisticated and efficient system of responses than those provided by instinctual automatism, could tend to question selfish approaches to human behaviour and rebalance the ongoing debate between viewpoints proposing human nature to be ontologically selfish or altruistic in a fundamental sense. Finally, this would open up the perspective to discuss and argue the existence of a shared axiological substrate among the human species, derived from our evolutionary history, amenable to scientific study, and open to consideration by an evolutionary Philosophy of biology philosophy that we have allowed ourselves to call scientific iusnaturalism, a natural law scientific theory.

## REFERENCES

- Beehner, J., & Kitchen, D. (2007). Factors affecting individual participation in group-level aggression among non-human primates. *Behaviour*, 144(12), 1551-1581. <https://doi.org/10.1163/156853907782512074>
- Benozio, A., & Diesendruck, G. (2015). Parochialism in preschool boys' resource allocation. *Evolution and Human Behavior*, 36(4), 256-264. <https://doi.org/10.1016/j.evolhumbehav.2014.12.002>
- Boesch, C., & Boesch-Achermann, H. (2000). *The chimpanzees of the Tai Forest: Behavioural ecology and evolution*. Oxford: Oxford University Press, USA. <https://doi.org/10.1093/oso/9780198505082.001.0001>
- Brewer, M. & Campbell, D. (1976). *Ethnocentrism and intergroup attitudes: East African evidence*.
- Brosnan, S. & De Waal, F. (2003). Monkeys rejects unequal pay. *Nature*, 425(6955), 297-299. <https://doi.org/10.1038/nature01963>
- Brosnan, S. et al. (2004). Fair refusal by capuchin monkeys. *Nature*, 428, 140. <https://doi.org/10.1038/428140b>
- Brosnan, S. et al. (2010). The interplay of cognition and cooperation. *Philosophical transactions of the Royal Society (B)*, 365(1553), 2699-2710. <https://doi.org/10.1098/rstb.2010.0154>
- Buchan, N., Grimalda, G., Wilson, R. et al. (2009). Globalization and human cooperation. *Proceedings of the National Academy of Sciences*, 106, 4138-4142. <https://doi.org/10.1073/pnas.0809520106>
- Choi, J. & Bowles, S. (2007). The coevolution of parochial altruism and war. *Science*, 318, 636-640. <https://doi.org/10.1126/science.1144237>
- Damasio, A. (2001). *El error de Descartes. La emoción, la razón y el cerebro humano*. Barcelona: Crítica.

- De Dreu, C. (2012). Oxytocin modulates cooperation within and competition between groups: an integrative review and research agenda. *Hormones and behavior*, 61(3), 419-428. <https://doi.org/10.1016/j.yhbeh.2011.12.009>
- De Dreu, C. et al. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science*, 328(5984), 1408-1411. <https://doi.org/10.1126/science.1189047>
- De Waal, F. (1997). *Bien natural*. Barcelona: Herder.
- De Waal, F. (2003). On the possibility of animal empathy. En T. Manstead, N. Frijda y A. Fisher (Comps.), *Feelings and emotions: The Amsterdam Symposium* (pp. 379-399). Cambridge: University Press.
- De Waal, F. (2007). *El mono que llevamos dentro*. Barcelona: Tusquets.
- De Waal, F. (2010). *Primates y filósofos. La evolución de la moral del simio al hombre*. Barcelona: Paidós.
- De Waal, F. (2015). *La edad de la empatía. Lecciones de la naturaleza para una sociedad más justa y solidaria*. Barcelona: Tusquets.
- De Waal, F. (2019). *El último abrazo. Las emociones de los animales y lo que nos cuentan de nosotros*. Barcelona: Tusquets.
- De Waal, F. (2022). *Diferentes. Lo que los primates nos enseñan sobre el género*. Barcelona: Tusquets.
- De Waal, F. & Roosemalen, A. (1979). Reconciliation and consolation among chimpanzees. *Behavioral Ecology and Sociobiology*, 5(1), 55-66. <https://www.jstor.org/stable/4599217>
- Descartes, R. (1991). *El tratado del hombre*. Madrid: Alianza.
- Diekhof, E., Wittmer, S., & Reimers, L. (2014). Does competition really bring out the worst? Testosterone, social distance and inter-male competition shape parochial altruism in human males. *PlosOne*, 9(7), e98977. <https://doi.org/10.1371/journal.pone.0098977>
- Dobzhansky, T. (1973). Nothing in Biology Makes Sense Except in the Light of Evolution. *The American Biology Teacher*, 35(3), 125-129. <https://doi.org/10.2307/4444260>
- Fearon, J. & Laitin, D. (1996). Explaining interethnic cooperation. *American Political Science Review*, 90, 715-735. <https://doi.org/10.2307/2945838>
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, 454(7208), 1079-1083. <https://doi.org/10.1038/nature07155>
- Fehr, E., Glätzle-Rützler, D., & Sutter, M. (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, 64, 369-383. <https://doi.org/10.1016/j.euroecorev.2013.09.006>

- Feigin, S. et al. (2014). Theories of human altruism: A systematic review. *Journal of Psychiatry and Brain Functions*, 1(1), 5. <https://doi.org/10.7243/2055-3447-1-5>
- Fry, D. (2018). The evolutionary logic of human peaceful behavior. En Verbeek, P. & Peters, B. (Eds.), *Peace Ethology. Behavioral Processes and Systems of Peace*. John Wiley & Sons Ltd., 249-265. <https://doi.org/10.1002/9781118922545.ch14>
- Fukuyama, F. (2001). Social capital, civil society and development. *2030 Third World Quarterly*, 22, 7-20. <https://doi.org/10.1080/713701144>
- Gazzaniga, M. (2012), ¿Qué nos hace humanos? *La explicación científica de nuestra singularidad como especie*. Barcelona: Espasa.
- Ghiselin, M. (1974), *The Economy of Nature and the Evolution of Sex*. Berkeley: University of California Press.
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge: Belknap Press of Harvard University Press.
- Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. London: Penguin.
- Hill, K. (2002). Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature: An Interdisciplinary Biosocial Perspective*, 13(1), 105. <https://doi.org/10.1007/s12110-002-1016-0>
- Hoffman, M. (1981). Is altruism part of human nature? *Journal of Personality and Social Psychology*, 40, 121–137.
- Hrdy, S. (2007). Evolutionary context of human development: The cooperative breeding model. En Catherine A. Salmon and Todd K. Shackelford (Eds.), *Family Relationships: An Evolutionary Perspective*. New York: Oxford Academic. <https://doi.org/10.1093/acprof:oso/9780195320510.003.0003>
- Hruschka, D. & Henrich, J. (2013). Economic and evolutionary hypotheses for cross-population variation in parochialism. *Frontiers in Human Neuroscience*, 7, 559. <https://doi.org/10.3389/fnhum.2013.00559>
- Hruschka, D. et al. (2014). Impartial institutions, pathogen stress and the expanding social network. *Human Nature*, 25, 567-579. <https://doi.org/10.1007/s12110-014-9217-0>
- Iacoboni, M. (2009). *Las neuronas espejo. Empatía, neuropolítica, autismo, imitación o de cómo entendemos a los otros*. Buenos Aires: Katz Editores. <https://doi.org/10.2307/j.ctvm7bdfw>
- Kustov, A. (2021). Borders of compassion: Immigration preferences and parochial altruism. *Comparative Political Studies*, 54(3-4), 445-481. <https://doi.org/10.1177/0010414020938087>



- Lemoine, S., Samuni, L., Crockford, C., & Wittig, R. (2022). Parochial cooperation in wild chimpanzees: a model to explain the evolution of parochial altruism. *Philosophical Transactions of the Royal Society B*, 377(1851), 20210149. <https://doi.org/10.1098/rstb.2021.0149>
- Levina, M. (2017). Under Lenin's watchful eye: Growing up in the former Soviet Union. *Surveillance & Society*, 15(3/4), 529-534. <https://doi.org/10.24908/ss.v15i3/4.6640>
- Lucchesi, S. et al. (2020). Beyond the group: how food, mates, and group size influence intergroup encounters in wild bonobos. *Behavioral Ecology*, 31(2), 519-532. <https://doi.org/10.1093/beheco/arz214>
- Mau, S., Mewes, J. & Zimmermann, A. (2008). Cosmopolitan attitudes through transnational social practices? *Global Networks*, 8, 1-24. <https://doi.org/10.1111/j.1471-0374.2008.00183.x>
- Morris-Drake, A., Christensen, C., Kern, J. & Radford, A. (2019). Experimental field evidence that out-group threats influence within-group behavior. *Behavioral Ecology*, 30(5), 1425-1435. <https://doi.org/10.1093/beheco/arz095>
- Mosser, A., & Packer, C. (2009). Group territoriality and the benefits of sociality in the African lion, *Panthera leo*. *Animal Behaviour*, 78(2), 359-370. <https://doi.org/10.1016/j.anbehav.2009.04.024>
- Parsons, K. et al. (2003). Male-Male Aggression Renders Bottlenose Dolphin (*Tursiops Truncatus*) Unconscious. *Aquatic Mammals*, 29(3), 360-362. <https://doi.org/10.1578/01675420360736532>
- Pisor, A. et al. (2020). Preferences and constraints: The value of economic games for studying *Human Behaviour*. *Royal Society Open Science*, 7, 192090. <https://doi.org/10.1098/rsos.192090>
- Radford, A. (2011). Preparing for battle? Potential intergroup conflict promotes current intragroup affiliation. *Biology Letters*, 7(1), 26-29. <https://doi.org/10.1098/rsbl.2010.0507>
- Radford, A., Majolo, B., & Aureli, F. (2016). Within-group behavioural consequences of between-group conflict: a prospective review. *Proceedings of the Royal Society B: Biological Sciences*, 283(1843), 20161567. <https://doi.org/10.1098/rspb.2016.1567>
- Reimers, L., & Diekhof, E. (2015). Testosterone is associated with cooperation during intergroup competition by enhancing parochial altruism. *Frontiers in Neuroscience*, 9, 183. <https://doi.org/10.3389/fnins.2015.00183>
- Rizzolatti, G. & Sinigaglia, C. (2006). *Las neuronas espejo. Los mecanismos de la empatía emocional*. Barcelona: Paidós.

- Samuni, L., Crockford, C., & Wittig, R. (2021). Group-level cooperation in chimpanzees is shaped by strong social ties. *Nature communications*, 12(1), 539. <https://doi.org/10.1038/s41467-020-20709-9>
- Samuni, L., Preis, A., Mundry, R., Deschner, T., Crockford, C., & Wittig, R. (2017). Oxytocin reactivity during intergroup conflict in wild chimpanzees. *Proceedings of the National Academy of Sciences*, 114(2), 268-273. <https://doi.org/10.1073/pnas.1616812114>
- Silva, A. & Mace, R. (2014). Cooperation and conflict: Field experiments in Northern Ireland. *Proceedings of the Royal Society B*, 281, 20141435. <https://doi.org/10.1098/rspb.2014.1435>
- Singer, P. (1981). *The expanding circle: Ethics, evolution, and moral progress*. Princeton: University Press.
- Ten Velden, F., Daughters, K., & De Dreu, C. (2017). Oxytocin promotes intuitive rather than deliberated cooperation with the in-group. *Hormones and behavior*, 92, 164-171. <https://doi.org/10.1016/j.yhbeh.2016.06.005>
- Thompson, F., Marshall, H., Vitikainen, E., & Cant, M. (2017). Causes and consequences of intergroup conflict in cooperative banded mongooses. *Animal Behaviour*, 126, 31-40. <https://doi.org/10.1016/j.anbehav.2017.01.017>
- Tokuyama, N., Sakamaki, T., & Furuichi, T. (2019). Inter-group aggressive interaction patterns indicate male mate defense and female cooperation across bonobo groups at Wamba, Democratic Republic of the Congo. *American Journal of Physical Anthropology*, 170(4), 535-550. <https://doi.org/10.1002/ajpa.23939>
- Wechkin, S. et al. (1964). Shock to a Conspecific as an Aversive Stimulus. *Psychonomic Science*, 1, 47-48. <https://doi.org/10.3758/BF03342783>
- Westermarck, E. (1912). *The Origin and Development of the Moral Ideas*. London: Macmillan. <https://doi.org/10.2307/1413437>
- Wilson, M., Hauser, M., & Wrangham, R. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, 61(6), 1203-1216. <https://doi.org/10.1006/anbe.2000.1706>
- Wilson, M., Hauser, M., & Wrangham, R. (2007). Chimpanzees (*Pan troglodytes*) modify grouping and vocal behaviour in response to location-specific risk. *Behaviour*, 144(12), 1621-1653. <https://doi.org/10.1163/156853907782512137>
- Wilson, M. et al. (2014). Lethal aggression in Pan is better explained by adaptive strategies than human impacts. *Nature*, 513(7518), 414-417. <https://doi.org/10.1038/nature13727>

- Wrangham, R. & Glowacki, L. (2012). Intergroup aggression in chimpanzees and war in nomadic hunter-gatherers: Evaluating the chimpanzee model. *Human nature*, 23, 5-29. <https://doi.org/10.1007/s12110-012-9132-1>
- Wrangham, R., Lundy, R., Crofoot, M., & Gilby, I. (2007). Use of overlap zones among group-living primates: a test of the risk hypothesis. *Behaviour*, 144(12), 1599-1619. <https://doi.org/10.1163/156853907782512092>



eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.31800>

# THE ANTHROPOCENTRIC BIAS IN ANIMAL COGNITION

## *El sesgo antropocéntrico en la cognición animal*

Carolina SCOTTO  
Universidad Nacional de Córdoba, CONICET  
carolinascotto@gmail.com

Recibido: 28/01/2024    Revisado: 15/02/2024    Aceptado: 01/03/2024

**ABSTRACT:** In the classical philosophical tradition, animals had the special function of serving as “objects of comparison” concerning humans. In that sense, philosophy adopted a peculiar comparative perspective focused on the categoric difference that separates humans from other creatures: an exceptionalist perspective. The Humanities developed an *anthropocentric canon* for the study of animals and privileged the search for differences over similarities of these with humans. On the other hand, the great boost that animal studies received under the influence of Darwin’s work promoted a different comparative perspective in the natural sciences. However, especially in comparative psychology, ingent efforts were devoted to avoid the errors that anthropomorphism would entail: attributing human properties to other creatures and privileging similarities over differences. It assumed that *anthropomorphic bias* entails a more fundamental type of error than *anthropocentric bias*. Now, this asymmetric diagnosis has been unmasked with different arguments. In the context of both disciplinary traditions, it is timely to reexamine the most persistent and negative manifestations of *anthropocentric bias* as a comparative bias for the study of animal cognition. In this work I will identify the following: the homogenization of animals into a single general category; psychological speciesism and the “de-mentalization” of

animals; the survival of a hierarchical conception of cognitive abilities; the selective application - only to animals - of Morgan's Canon or *anthropodenial* and its complement, the assumption of idealized mental capacities in the human case or *anthropofabulation*; asymmetrical or distorsive methodological strategies for the study of animals versus humans which affects the comparative interpretations; and different manifestations of semantic anthropocentrism.

*Keywords:* animal cognition, anthropocentric bias, anthropomorphic bias, anthropodenial, anthropofabulation.

**RESUMEN:** En la tradición filosófica clásica, los animales tuvieron la función especial de servir como "objetos de comparación" con respecto a los humanos. En ese sentido, la filosofía adoptó una peculiar perspectiva comparada, centrada en la diferencia categórica que separa a los humanos de otras criaturas: una perspectiva excepcionalista. Así, las Humanidades elaboraron un *canon antropocéntrico* para estudiar a los animales y privilegiaron la búsqueda de las diferencias sobre las semejanzas de estos con los humanos. Por su parte, el gran impulso que recibieron los estudios animales bajo el influjo de la obra de Darwin, alentó una perspectiva comparada diferente en las ciencias naturales. Sin embargo, sobre todo en la psicología comparada, se dedicaron ingentes esfuerzos a evitar los errores que conllevaría el antropomorfismo: atribuir propiedades humanas a otras criaturas, privilegiando las semejanzas antes que las diferencias. Se asumió con frecuencia que el *sesgo antropomórfico* acarrea un tipo de error más fundamental que el *sesgo antropocéntrico*. Ahora bien, ese diagnóstico asimétrico ha sido desmentado con distintos argumentos. En el contexto de ambas tradiciones disciplinares, resulta oportuno reexaminar las manifestaciones más persistentes y negativas del *sesgo antropocéntrico* como un sesgo comparativo para el estudio de la cognición animal. En este trabajo identificaré las siguientes: la homogeneización de los animales en una única categoría general; el especismo psicológico y la "desmentalización" de los animales; la supervivencia de una concepción jerárquica de las capacidades cognitivas; la aplicación selectiva -sólo a animales- del Canon de Morgan o *antroponegación* y su complementaria, la suposición de capacidades mentales idealizadas en el caso humano o *anthropofabulación*; las estrategias metodológicas asimétricas o distorsivas para el estudio de animales vs humanos que afectan las interpretaciones comparadas; y distintas manifestaciones del antropocentrismo semántico.

*Palabras clave:* cognición animal, sesgo antropocéntrico, sesgo antropomórfico, antroponegación, antropofabulación.

## 1. INTRODUCTION

“...man is godlike, animals thinglike...[but] reason looks to me suspiciously like the being of human thought; worse than that, like the being of one tendency in human thought.”

(Coetzee, *Elizabeth Costello*, p. 23)

In much of the Western philosophical tradition, for example, in the works of Aristotle, Descartes, and Kant, non-human animals<sup>1</sup> had the special function of serving as *objects of comparison* concerning humans (Glock, 2012). It thought the search for a specific difference would be strengthened by “looking down”. Thus, animals are represented as the closest contrast to human “nature.” Condillac summarized the motivation for this perspective: “It would be of little interest to know what animals are if it were not a means of knowing what we are”.<sup>2</sup> In that sense, the great philosophical tradition adopted a peculiar comparative perspective, based on a categorical difference that separate humans from other creatures: an exceptionalist perspective.

The Humanities and the different “humanisms”, including some sophisticated elaborations of the recent philosophy of the animal mind, built an *anthropocentric canon* that extended to all areas of culture. They laid the foundations for the exclusion of animals from Humanities first, and Social Sciences later. According to the canon, animals “belong” only to the natural sciences. For its part, natural-scientific studies on animals, in disciplines such as ethology and comparative psychology, even after Darwin, had to fight an arduous battle against the *anthropocentric canon* for much of the last century. However, the epistemological and methodological discussions were dominated by the “spell of anti-anthropomorphism” (de Waal, 2003), more than by avoiding anthropocentrism.

*Anthropocentrism* is usually defined as the view according to which cognitive characteristics assumed to be distinctive or unique to human beings are taken as a reference or standard, explicitly or implicitly, for studying and evaluating the cognitive abilities of animals. In this paper, I will only refer to anthropocentrism about cognitive, psychological, or

1. I will refer interchangeably to *animals* or *non-human animals*. But, as will be seen, these and other ways of referring to all species, excluding humans, are not always neutral.

2. E. de Condillac (1755), *Traité des Animaux*, (Paris: Vrin, 1987), 1. (Quoted and translated by Glock, 2012).

mental abilities (I will use these terms interchangeably<sup>3</sup>). In this context, cognition refers to the processes or mechanisms through which a creature perceives, learns, and categorizes, which allows it to have flexible and goal-oriented behaviors (Andrews & Monsó, 2021).

I will try to show that anthropocentrism adopts the structure of a *comparative bias* that continues to influence, through certain persistent but mostly unnoticed tendencies in ways of thinking about animal minds. This bias also influences our general view of animals and the way we conceive our epistemic, moral, and even political relations with them. My purpose is to identify some indicators that evidence the survival of anthropocentrism in scientific and philosophical studies on animal cognition.

de Waal (2003) observes that "...the way we look at other animals reflects the way we look at ourselves...", not only in the natural sciences (which he refers to) but also in the human sciences. Thus, anthropocentrism is usually presented as a fairly natural perspective, rather than a distorted view of animals. That is why it is interesting to examine it as a bias, or a set of biases, rather than as an approach or a conception. As a bias, it has been more debated in the ethical and political literature on animal rights than in the literature on animal cognition. In other words, it seems essential to reflect on how certain characteristic forms of self-understanding permeate our view of other creatures, generating an anthropocentric comparative variety of the Socratic injunction "Know thyself," which distorts our approach to animals. Encouraging an approach free of this anthropocentric bias could indirectly contribute to dismantling the unjustified and persistent gaps between the human and social disciplines and the natural science disciplines of animal studies based on this bias.

As many authors point out, many dimensions are involved in the question of whether animals have some mental capacities, and how it is possible to establish them justifiably. I will refer to some of them: ontological, epistemological, methodological, and semantic. Even though they are closely related and sometimes even strongly interpenetrated - because a question may formulate at one level but obliquely involve or point to another - all must be taken into account for an adequate understanding of the effects of *anthropocentrism* in studies on animal cognition.

3. *Anthropocentrism* is usually defined in terms of different human characteristics, not only cognitive but also agential and moral. The term also refers, more broadly, to a sort of relationship with the environment or nature, including animals, intelligent machines or artifacts, and even hypothetical suprahuman entities.



## 2. ANIMALS AS “OBJECTS OF COMPARISON”: THE ANTHROPOLOGICAL DIFFERENCE

The concept of *anthropological difference* developed in continental philosophy, refers to the qualities or characteristics that would make the human species unique, i.e., different from any other. Such a difference should explain, in turn, other differences derived from it (cf. Glock, 2012). It would be one or more traits universally shared by all humans, but only by them, that is, traits that distinguish them as essential or categorically different creatures. In other words, it must be a fundamental difference, in the sense that it can explain our *unique* status (cognitive, agential, moral): a *specific difference*, in the strictest sense of the word. This is an ontological variety of anthropocentrism.

Whether or not such a difference exists, whether it is a qualitative one, or consists of a chain of traits similar to those of other species that combined produce, at last, a qualitative leap, or whether it is only a difference of degree, remains a debatable issue. Conjectures about explaining human cognitive uniqueness have reached sophisticated formulations in recent years and the phenomenon has even been described as an “epidemic” (Shettleworth (2012). Now, the interest in establishing what the differences are between species concerning a trait, whether cognitive or not, is only a way of approaching the unobjectionable purpose of explaining biological diversity and understanding the phylogenesis of some characteristics and abilities. But then, why do similar theoretical motivations not lead specialists to wonder about the uniqueness of the jaguar, the condor, or any other species?

On the other hand, if the question of human *uniqueness* is motivated by the purpose of establishing what abilities dramatically differentiate us from all other species, that explains why many of these attempts adopt a circular format: if research reveals that a candidate trait is not ultimately unique, instead of continuing to investigate its distribution in different species, the search begins for another trait that can satisfy the claim of *uniqueness* (cf. Vasilieva, 2019). This motivation has frequently been driven by the advances in knowledge of many species to the point that they have defeated or called into question proposed demarcation lines. Furthermore, if the comparison was intended to justify human cognitive *uniqueness*, it would also have to be done with all other extant species, which is impossible. So, “theories of human uniqueness are inevitably provisional” (Shettleworth, 2012, p. 2795). On the other hand, this type of search is often guided by the assumption that a given trait is categorically present

or not, and even by adherence to the progressive view of the *Scala Naturae*, in contrast to a gradualist conceptualization. At the same time, it has not always been taken into account that a trait might not be specific to a species, that is, be present in more than one species, but only be *typical* or *habitual* in one of them, and *occasional* in others. Thus, not only the uniqueness but also the distribution of a trait in different species could shed light on the evolution of said trait (cf. Vasilieva, 2019). In other words, the evolutionary significance of a given trait can be better understood when comparative research abandons the categorical, dichotomous, and anthropocentric presuppositions that are associated with the search for “unique” human traits, that is, *anthropological difference*.

In philosophical and, in general, humanistic literature, comparative reflection between humans and animals was clearly dominated by the tendency to establish cognitive *differences* between humans and other species. Thus, the human and social sciences were constituted as “sciences of discontinuity” (Noske, 1993) by assuming a clear separation between them and the natural sciences with regard to the studies on humans and animals, respectively. This is a dimension in which anthropocentrism even adopted the rank of the constitutive norm of humanistic disciplines (cf. Suárez-Ruiz, 2021). Now, it might be thought that this epistemological discontinuity reflects a stage already surpassed. However, varieties of discontinuism and anthropocentrism remain the default assumptions in the most vigorous theoretical traditions in the social and human disciplines. Next, I will make a brief reference to some of them.

It is known as the “animal turn” (Ritvo, 2007) to a new interdisciplinary field of socio-scientific and humanistic studies (linked to continental philosophy) whose focus of interest is “human-animal relations”, not animals. In these approaches, “human centrism” expands the circle to encompass animals, but on condition of turning them into social constructions, more specifically, into *discursive creatures*. Thus, abundant “post” declamations that concern animals, delivered from some non-anthropocentric counter-pinnacle, do not seem to have overcome the barriers between the natural and the “uniquely human.” In some versions of this “turn”, they even try to subsume the natural into the human (sic). Thus, despite strongly questioning the classical anthropocentric tradition, they can hardly offer anything more than new ways of conceptualizing the different relationships of dependence of animals with the human center of reference. In other tradition, naturalized ways of philosophizing are edifying an empirically informed philosophy of animal minds. But, as will be seen below, they coexist with influential expressions of Western canonical thought.

On the other hand, natural scientific studies on animal cognition, especially in ethology and psychology, encouraged a different comparative perspective: identifying similarities and differences in psychological abilities between species on an evolutionary *continuum*. In its scope, “the default assumption is continuity” (de Waal, 2016, p. 124). The canonical expression of the continuity thesis is due to Darwin: “...the difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind” (1871, p. 105). However, comparative psychology has presented a more explicit battlefield in which more or less “continuist” or “discontinuist” hypotheses were debated. (Vasilieva, 2019). Comparative psychology, unlike ethology, inherited the essential unjustified distinction between humans and other animals, long assuming that the species with which any other species had to be compared was the human species. Or even that, given the goal of explaining the evolution of human cognition, only the “most intelligent” animals should be studied. (Döring & Chittka, 2011; Chittka et al., 2015). Again, the study of animals was motivated by the interest in understanding human behavior and mind. Now, it is manifestly objectionable from an evolutionary point of view to intend to understand the cognitive similarities and differences between species by taking as paradigm human capacities and afterward by grouping animals to encourage comparison with them. In other words, what abilities to investigate in non-human animals and how to study them should not depend on human cognitive adaptations or how they are studied. On the contrary, it would be better to investigate *species-neutral* cognitive abilities (Figdor, 2021). Namely, we must compare the cognitive abilities of different species with each other, not only with those evolutionarily close or in which abilities derived from processes of convergent evolution are evident. It must explain how cognition evolves, whether it be through phylogenetic evolution, convergence, or homology (Chittka et al., 2015). In any case, “...humans are not necessarily central to every comparison” (de Waal, 2016, p. 28).

de Waal & Ferrari (2010) identified the contrast between the predominant perspective in comparative psychology, which they call *top-down*, and the inverse *bottom-up* perspective that propose in its replacement. The former focuses “on the pinnacle of cognition,” asks “all-or-nothing questions,” and has “an obsession with rankings and the human-animal divide.” Species are distinguished by clear dividing lines and traits studied are characterized as “unified” capacities rather than “multicomponent structures.” Instead, the *bottom-up* perspective focuses on “the basic building-blocks of cognition [which] might be shared across a wide range of species” (2010, p. 201). Under this evolutionary perspective it is possible to identify different combinations of traits partially shared between

some species, performing functions that may also differ. So, they propose unifying ethology with comparative psychology in the new discipline of *Evolutionary Cognition* dedicated to the “study of all cognition, human and animal, from an evolutionary perspective.” This new approach recommends exploring “common denominators first before exploring species-typical specializations” (2010, p. 205). Moreover, to obtain “...a unitary theory that covers all the various cognitions found in nature...[T]o create space for this project, I recommend placing a moratorium on human uniqueness claims” (de Waal, 2016, p. 158). This way, it would be easier to discourage the human-centered perspective. Let’s see what kind of bias anthropocentrism is.

### 3. ANTHROPOCENTRISM AS A BIAS

Despite what the inflection suggests, *anthropocentrism* does not always refer to an explicit and elaborate approach. It can also be conceptualized as a bias (Buckner, 2013; Andrews & Monsó, 2021; Andrews, 2020b). A bias is defined as the tendency or inclination, more or less systematic, to form prejudiced, flawed, or distorted judgments or evaluations, often implicit or unconscious, e.g. difficult to put under reflective control (Kahneman, Slovic & Tversky, 1982; Holroyd & Sweetman, 2016). This tendency or inclination affects different dimensions related to the study of a certain phenomenon: the concepts used, the questions and hypotheses that can be formulated, the admissible methods, the nature of the evidence, and the explanatory power of the models or theories to which they apply. By the way, the effects of such biased judgments or evaluations affect the most varied areas and are not only theoretical.

I call *anthropocentric bias* the tendency to consider that human characteristic ways of experiencing, conceiving, and thinking provide the criterion or “gold standard” for understanding the behavior of non-human animals. To the extent that the human standard is conceived in a biased manner, for example, as an overly intellectualized conception of human mental abilities, anthropocentrism will imply a high degree of contrast between humans and other animals. Although the expression *anthropocentric bias* has some use in the literature that analyzes the topics we are interested in, it is not clearly defined or used systematically.

Many biases take the form of “centrism” in the sense of “human-centered”: ethnocentrism, androcentrism, etc., which means that a human group with some characteristics associated with it provides the norm or

standard for evaluating others. Those who locates in the center embody “the indisputable value”, around which an “expanding circle” opens (Andrews, 2020a). Therefore, anthropocentrism is a trend that leads to “a human-centered approach.” About animals, it consists of “[holding] the human mind [to be] the gold standard against which other minds must be judged” (Povinelli, 2004, p. 29).

In this paper, I will understand both anthropocentrism (and its varieties) as well as its counterpart, anthropomorphism, as *comparative biases* (Buckner, 2013, 2023), as it affects comparisons between humans and animals. Regarding the latter, I will try to show that, despite the unequal attention it has received in contrast to the anthropocentric bias (Buckner, 2013), “anthropomorphism and anthropocentrism are never far apart: the first is partly a “problem” due to the second” (de Waal, 1999, p. 256). It should also be noted that the characterization of anthropomorphism as a bias is very widespread. As such, it is explicitly questioned, especially in comparative psychology. It contrasts with anthropocentrism, which rather is treated as a doctrine or a body of ideas. In summary, *anthropocentric bias*, particularly its role in studies on animal cognition, has been underexplored<sup>4</sup>.

In a strong or substantive sense, *anthropocentrism* assumes a categorical distinction between humans and other creatures and, on that basis, holds that humans provide the standard for judging non-humans (Andrews, 2020a). As we will see later, our language is permeated by anthropocentric categorical distinctions, all of which generate a “distancing” effect on human centrality. The concept has received different labels: *inverted anthropomorphism* (Romanes, 1882), *anthropodenial* (de Waal, 1999), *anthropectomy* (Andrews & Huss, 2014). These notions include different nuances. Note that the first is formulated taking as reference its opposite: *anthropomorphism*. If this last notion refers to the (unjustified) tendency to attribute human properties or characteristics to other entities, the opposite tendency would lead to differentiating humans, and identifying their singularities. However, the concept seems to capture another idea: human differences are especially relevant.

As such, anthropocentrism is a variety of *speciesism* understood as “...a prejudice or attitude of bias in favour of the interests of members of one’s own species and against those of members of other species’ (Singer, 2009[1975], 6). Although this definition is species-neutral, it is

4. Andrews’s (2020b) book is an exception because it is dedicated to the detailed analysis of the different biases that affect studies of animal cognition in comparative psychology, including a special treatment of anthropocentric bias.

generally used to refer to the human species and, by extension, a group within it. As is known, the term was coined by analogy with others such as *sexism* or *racism*, to refer to the comparative bias that consists of the “unjustified consideration or treatment of those who are not classified as belonging to a certain species” (Horta & Albersmeir, 2020). Figdor (2021) calls *psychological speciesism* the conception according to which the possession by humans of some “superior” cognitive abilities on a hierarchical scale is the basis of the recognition of the superior value of human life compared to that of other species, *moral speciesism*. Many humanists views adopted both conceptions. Psychological speciesism is the same phenomenon that I call anthropocentric bias as it applies to animal cognition: human cognition is the standard or criterion implicit against which it is established whether and to what degree non-human creatures possesses some capacity. As we will see in the next section, the homogenization of animals is the first step toward a restrictive approach to animal cognition.

#### 4. HOMOGENIZATION OF ANIMALS

A first manifestation of the anthropocentric bias, and “...the decisive step... that seemed quite innocent to us” (Wittgenstein, 2009, IF §308), it is the generic reference to animals, trying to first *assimilate* them and then *contrast* them with humans. I am not referring to the use of one generic label or another as a shorthand, that might be justified on grounds of expository economy, but to the (usually implicit) suggestion that the same questions are appropriate concerning such a heterogeneous set of species, all except humans: Do they have concepts? Can they reason? Do they understand “other minds”? These questions, for their part, are those considered relevant for studying the human mind. Thus, this assimilation is based on the previous adoption of the anthropocentric standard on our way of conceptualizing animals. Besides, we are imposing it *a priori*, that is, to establish the preconditions for comparative research. This assumption is grounded on a previous one: all non-human animals are more similar to each other than some species are to our own. Now, it is wrong to presuppose that all non-human species possess sufficient and relevant common characteristics, cognitive or otherwise, so that the use of a single label is justified. Finally, as noted, “animal” (and the like) does not refer to anything that can be studied scientifically.

A classic version of this dichotomous view assumed that all animal behaviors could be explained by basic mechanisms of the same type.

Consequently, any animal could serve to establish the desired contrast with humans: experiments with pigeons or rats could be sufficient to provide valid evidence for all species. As de Waal points out, Skinner's theory outlined in *The Behavior of Organisms* exemplified this idea in an extreme way (de Waal, 2016, p. 27). On the other hand, in the post-behaviorist era, it is often assumed that the very concept of cognition (or mind) is instantiated by human cognition or, in other words, that there is "a paradigmatic cognition": human cognition (Allen, 2017; Vasilieva 2019; Bräuer et al., 2020). As part of the same bias, especially in the philosophical literature, it is common to distinguish between cognition *strictu sensu* and other not fully cognitive capacities (e.g., associative, perceptual, non-representational, non-intentional, "non-genuine", etc.) (cf. Allen, 2017). These contrasts assume that *all animals* differs from humans. But, just as there is a diversity of species, we should speak in plural of "cognitions" or "intelligences", and even more so, of abilities and performances, with different characteristics in different species (de Waal, 2016, p. 12).

If, otherwise, only a single shared negative characteristic is presupposed – that is, all animal species would not possess some human cognitive capacity, or would not possess it to the same degree and/or in the same way, how can we avoid the objection that it is, once again, a capricious assimilation of all of them? After all, we could group some animal species according to different criteria, choosing any negative characteristic (or a group of them) to justify different classifications, e.g., one could include some animal species and the human species. Furthermore, if the characteristic we are interested in identifying is not present in some species but is present in the human species, what comparative conclusions can we expect to draw? Could they not also be obtained from similar comparisons between other species, identifying other unique traits in some of them? As for the cognitive differences that could be found, don't they have the same relevance as the distinctive characteristics of some animal species, for example, echolocation in bats, that are not part of the human repertoire? In short, we know that "[T]here are lots of wonderful cognitive adaptations out there that we don't have or need. This is why ranking cognition on a single dimension is a pointless exercise" (de Waal, 2016, p. 12).

The homogenization of animals is not avoided by sharing the term *animals* to refer differently to *human animals* and *non-human animals*. Because, beyond the double use of the term (that is, we are *all animals*), these denominations not only function like the classic distinction "humans versus animals" as a dichotomy but, in addition, they explicitly attribute to all *non-humans* animals a negative property defined in reference to

humans. This term "...groups millions of species by an absence, as if they were missing something. Poor things, they are not human!" (de Waal, 2016, pp. 27-28). The anthropocentric bias of this label is evident. de Waal suggests, sarcastically, that to be consistent with this terminological choice, whenever we want to refer to a particular species, we should also specify that it is also not all the others: nonpigeon, nonpenguin, nonelephant, and so on (2016, p. 28).

This terminology, furthermore, induce us to characterize animals as creatures that lack something: homogenization based on a negative property also means attributing to them a certain inability. Wittgenstein highlighted this meaning of *deprivation* of the anthropocentric bias in relation to linguistic ability, and outlined a critique:

It is sometimes said: animals do not talk because they lack the mental abilities. And this means: "They do not think, and that is why they do not talk". But they simply do not talk. Or better: they do not use language -if we disregard the most primitive forms of language. (PI §25) (I underline).

That is, although certain "forms of language" are present in some species but not in others, what sense would it make to say that the latter are missing something? It makes as little sense as claiming that humans lack, for example, the ability to communicate through the amazing vocal and visual displays that many birds use. We simply do not dance or sing like them.

Different labels have been chosen in philosophical literature to express explicitly this meaning of *deprivation*, that is, to define animals as lacking some capacity. Some of them add a derogatory connotation: "thoughtless brutes" (Malcolm, 1972-73), "languageless creatures or "dumb creatures" (Davidson, 1985), "speechless brutes" (Heil, 1982), "mere animals", (Mc Dowell, 1996), "non-rational animals" (Mc Dowell, 2009), and "non-discursive creatures" (Brandom, 2000), among many others. As can be seen, these categories oscillate between characterizing animals as *mindless* or, in more recent versions, *mentally limited*, or *lacking one or another capacity* considered critical or unique, mainly language. Now, these labels can only make sense in the Aristotelian conception of the *Scala Naturae*, which places all non-human species in the same negative place in the hierarchy of beings, rather than in a darwinian gradualist but not progressionist view of evolutionary continuity.

In particular, the terms "brutes" and "beasts", frequently used by philosophers of different eras, although they would have had a more or less neutral meaning in the classical philosophical literature, gradually began to acquire a more derogatory meaning (cf. Preece, 2005). Surprisingly, so



many learned philosophers have continued employing this terminology even in this century. The concept of “dumb beasts,” which attempts to capture the classical conceptualizations of animals by many ancient philosophers (cf. Osborne, 2007), would be appropriate to cover the positions of many conspicuous representatives of recent analytical philosophy, such as the authors already mentioned.<sup>5</sup> This is because they assume not only a hierarchical and exceptionalist perspective but also one imbued with another particular bias: they give language critical importance in their arguments in favor of human uniqueness (which, curiously, is proportional to what they give to the object primary of its disciplinary *métier*). Among the most prominent contemporary philosophers, Davidson developed a canonical formulation of the constitutive link between cognition and language:

One belief demands many beliefs, and beliefs demand other basic attitudes such as intentions, desires, and if I am right, the gift of tongues ... To make the distinction so strong, and to make it depend on language, invites an accusation of anthropocentrism. The complaint is just, but it ought not to be leveled against me. I merely describe a feature of certain concepts. ... We connive with our language to make it, and us, seem special (1985, pp. 318-319).

In any case, Davidson makes it clear that he is not interested in examining whether this or that species has certain capacities or not, but only in addressing the question of the nature of (human) thought by adopting this comparative approach “*as just as a colorful...way of thinking...*” of it. (1985, p. 319). (I underline). But if anthropocentrism is a characteristic of “our concepts,” Davidson is suggesting that “our concepts”, that is, “our language”, could not be anything else and, consequently, anthropocentrism is simply unquestionable. Similarly, Brandom justifies the contrast between discursive and non-discursive creatures by appealing to his own theoretical interests: “I am more interested in what separates concept users from non-concept users than in what unites them” (2000, p. 3). He simply “makes explicit” that his project prioritizes the “distinctive” or even the “exceptional,” vgr., “discontinuities” instead of “continuities.” But that “priority” is not innocent because produces a distorting view on the magnitude of the contrast: two types of creatures. Mc Dowell, for his part, distinguishes categorically between rational and non-rational animals or “the

5. Malcolm epitomizes this position: “...the relationship between language and thought must be so close that it is really senseless to conjecture that people may not have thoughts, and also senseless to conjecture that animals may have thoughts” (1973, p. 17).

rest of the animal kingdom”, assuming the traditional distinction between rational abilities or “responsiveness to reasons *as such*” and behavior “determined by nature”, respectively (Mc Dowell, 2009, p. 128).

Arguments based on human linguistic or discursive competencies received a variety of idealized formulations, creating an articulated and expanding circle or “space of reasons” inhabited only by human minds: concepts require concepts that combine in propositions, propositions give their content to mental states, basically beliefs, first-order beliefs require second-order beliefs, all these states are expressed paradigmatically through statements linked together by inferential links, etc. Finally, one comes to the conclusion that “...the distinction between human and non-human depends on whether or not the difference between a subject and a predicate is known” (Coetzee, 1999, p. 66). But doesn't this way of understanding “mentality” lead to the conclusion that animals do not speak and then do not think because they could not become philosophers of mind, language, etc.? I am here paraphrasing Flack and de Waal (2000) when they state that, although they are moral creatures, “animals are not moral philosophers.” (p. 23).

Under the imprint of Davidson's style arguments, many other philosophers developed refined varieties of epistemic skepticism about the *dumb minds of animals*: for example, presupposing that linguistic evidence is necessary for the attribution of mental states to others, i.e., not admitting that, in the human case, it can only be sufficient and only under certain conditions. (Andrews, 2020b); by assuming that the standard of “propositional precision” is provided by human thought and language (under a homogenizing characterization of both)<sup>6</sup>. In any case, there is not much distance between these and Cartesian ideas:

...the reason why animals do not speak as we do is not that they lack the organs but that they have no thoughts. It cannot be said that they speak to each other but we cannot understand.” (Descartes, 1991, p. 303).

6. Dennett (1998) synthetically formulates the problems that lead to what he calls “the misguided goal of propositional precision”. It applies to humans and animals when *expression* and *description* of mental contents are confused. Human language, even each particular language, imposes a format that makes it difficult to *express* what speakers of different languages think and virtually impossible what animals think. But *describing* those contents is perfectly possible, although it may be difficult. The kind of precision made possible by the linguistic format is not always appropriate to the discriminatory power of the animal contents. But human language is not a “universal measurement system” of mental contents. In short, we can... exhaustively describe what we cannot express, without leaving any mysterious residue at all”. (p. 42)

But “All human beings use [it] (language) however stupid and insane they may be, even though they may have no tongue and organs of voice; but no animals do. Consequently this can be taken as a real specific difference between humans and animals.” (Descartes, 1991, p. 366)

Thus, *brutes* have no (our) language because they have no (our) reason. ¿How should we interpret such kind of “conclusions”?

## 5. “HIGHER” AND “LOWER”?

One might believe that evolutionary thought managed to propagate the conceptualization and language appropriate to the approach to species differentiation (“common descent with modification”). Or what is the same, that the idea of a *Scala Naturae* or *Great Chain of Being* (Lovejoy, 1936), understood as a scale of progress, is currently only a pre-scientific concept in disuse. This idea is structurally linked to anthropocentrism because that *Scala* has the human species at the top as the most evolved or “highest” and not just another branch of the tree of life. It is, therefore, a central component of the anthropocentric bias. Darwin himself reminded himself that he should avoid this terminology: “Never use the words higher and lower”.<sup>7</sup> On another occasion, he seems to admit both their power of persuasion and the error to which these concepts lead:

With respect to ‘highness’ and ‘lowness’... It appears to me that an unavoidable wish to compare all animals with men, as supreme, causes some confusion; and I think that nothing besides some such vague comparison is intended, or perhaps is even possible, when the question is whether two kingdoms such as the articulata or mollusca are the highest. (I underline)<sup>8</sup>

The notable fact is that this conceptualization is not absent in recent specialized literature. de Waal points out that in comparative psychology, the “habit of dividing animals into “higher” and “lower” forms has persisted for a long time...” viewing animals “...as mere stand-ins for humans: a monkey is a simplified human, a rat a simplified monkey, and so on” (2016, p. 27). A study conducted on articles published in prestigious academic journals in evolutionary biology over the past decade shows that

7. Darwin (1845), note written on the margin of his copy of Robert Chambers’ *Vestiges of The Natural History of Creation* (1844). Cited by Mayr (1988), p. 251.

8. Charles Darwin (1854), “Letter to Joseph Hooker”.

“progressionist language” or a “pre-evolutionary language” is still very common.<sup>9</sup> Authors conclude that “...the great chain of being is still with us, 153 years after Darwin (1859) published *The Origin of Species*, eventually paving the way to modern tree-thinking...” (Rigato & Minelli, 2013, p. 5). A similar research by Ullrich, Mittelbach & Liebal (2017) proposed to identify both explicit and implicit indicators of “norms of progress” or vestiges of them, a bias that assumes linear progress of evolution from primitive to improved traits or organisms and which adopts progressionist categories. The research was based on a corpus of scientific publications on “language” and “communication”, and shows a range of “higher” and “lower” categories associated with humans and non-human primates and other species, respectively.<sup>10</sup>

The terminology “higher” and “lower” (which is currently usually used in quotation marks to suggest a meaning that is not strictly literal but is nevertheless desired to be preserved) survives in other notions, also graded and with equivalent meanings, concerning to cognitive abilities: “more simple”, “primitive” versus “more sophisticated”, “advanced”, “complex”, often conceived as “more evolved” or “superior” in the sense of a “ladder of progress.” It would be interesting to explore the connotations of all these concepts. I can only point out here that their use is associated, in many cases, with anthropocentric standards: they presuppose the presence or absence, respectively, of agency, that is, autonomy, rational deliberation, flexibility, and self-awareness characteristically human.

Now, once the idea of the *Scala Naturae* has been questioned and beyond the debate between continuists and discontinuists, and even admitting that the human species could be cognitively unique (as all species are, on the other hand), whether in kind or degree (this point is also not relevant here), the hierarchical perspective on species and their abilities is unjustifiable. But then, how could we consider that one of them, the human, can provide the standard or criterion for evaluating others? Besides, the links of continuity between related species do not go in a single direction (Andrews, 2020b). Thus,

Any human-like capacity in a non-human species is also a non-human-like capacity in humans, but the symmetry of similarity is broken by this criterial status. It follows that cognitive capacities must be defined in

9. The study covers 16 prestigious specialized journals on more than 67,000 papers published between 2005 and 2010. The results yielded 1,287 expressions that fit into the language of the *Scala Naturae* (“lower X” or “higher X” applied to organisms and species), including general journals such as *Science* and *Nature*.

10. The study included 915 journal articles published between 2005-2015.

ways that do not take human cognitive capacities or features of them as criterial in comparative research. (Figdor, 2021)

Moreover, this same hierarchical idea is what allowed for many centuries to adopt extreme positions or attitudes of “dementalization”, that is, “depsychologize” animals. These attitudes are currently manifested more covertly, in different contexts. For example, denying mental abilities to animals facilitates many practices about them, such as protecting the cultural practice of eating meat by minimizing its negative effect on the animals involved (see Bastian et al., 2012). *Epistemic speciesism* is another expression of the same dementalization operation. It is expressed in efforts to purge the sciences of animal cognition of anthropomorphic “risk,” as we will see in the next section.

## 6. BEHAVIORISM FOR ANIMALS, INTELLECTUALISM FOR HUMANS<sup>11</sup>

A feature indicative of the presence of an *anthropocentric bias* reflected in the widespread adoption of the Morgan Canon (MC) as a presupposition guiding standard practice in comparative psychology. As we know, MC was proposed as a corrective to *anthropomorphic bias*, but its justification and effects deepened *anthropocentric bias*. As we have said before both biases are closely linked: “anthropomorphism and anthropocentrism are never far apart: the first is partly a “problem” due to the second” (de Waal, 1999, p. 256). The latter reflects in the tendency to the selective application of MC only to animals. It is combined with the assumption of idealized mental capacities as typical of the human case or the *anthropofabulation bias* (Buckner, 2013). I will briefly refer to each of these phenomena. Let us first remember that the MC recommends that:

In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale.” (Morgan, 1903, p. 292)

It is not my intention to discuss this epistemic principle here. There are excellent assessments in the literature about its possible interpretations, from more prohibitive to merely precautionary versions (see Fitzpatrick,

11. This formula inherits Nozick’s scheme about the moral status of human and non-human animals: *utilitarianism for animals, Kantianism for people*, and it is an alternative version of *behaviorism for animals, representationalism for people* (cf. Borchert & Dewey, 2023).

2018); about its proper understanding in the context of Morgan's work; and about its positive and negative impact on comparative psychology (de Waal, 1999; Sober, 2005; Fitzpatrick, 2008; 2018; Buckner, 2013; Andrews and Huss, 2014; Andrews, 2020). To my objectives, I will limit myself to pointing out under what interpretations it became, directly or indirectly, a tool of *anthropocentric bias*. As part of the orthodoxy of the discipline, there is no doubt that he favored epistemic speciesism: either animals lack minds or have fewer or "lower" mental capacities than humans. This is because it stimulated not merely skepticism or agnosticism, but denialism about them.

We leave aside the distinction between "higher" and "lower" in the formulation of the MC because it considered in the previous section. However, alternative conceptualizations are equally objectionable since they presuppose a categorical and hierarchical distinction between "higher" and "lower" cognitive mechanisms. These are problematic distinctions, even according to their ecumenical interpretations (see Buckner, 2017), at least when these mechanisms are considered at different levels of abstraction (Andrews, 2020b). Indeed, not only would it be assumed that there are no in-between cognitive abilities and mechanisms (Buckner, 2013), but it would also be denying that different types of processes (associative and cognitive) may be necessary to explain the varied behaviors of species, including *Homo Sapiens* (Andrews, 2009; Buckner, 2017).

MC as an epistemic and methodological principle is also presented as the standard of scientific rigor in animal psychology, since it would allow us to avoid the effects of anthropomorphic bias: the unjustified attribution of human capabilities to other creatures (Wynne, 2004). However, the anti-anthropomorphic concern is revealing the same assumption already noted: the existence of a hierarchical scale from lowest to highest cognitive complexity, whose pinnacle is the human mind. On the other hand, MC has served to justify an "a priori resistance" to admitting mentalistic explanations of animal behavior. So it justifies conceiving as the expression of a bias, that is to say, "a persistent impediment to progress" (Fitzpatrick, 2008, p. 225) or "a pre-empirical obstacle" to research (Andrews, 2009, p. 52). De Waal calls *anthropodenial* "the a priori rejection of characteristics shared between humans and animals when in reality they may exist" (1999, p. 258). With this name he seeks to highlight the biased character of anti-anthropomorphism, thus discrediting its supposed status as an epistemic principle that seeks to ensure scientific rigor and objectivity. Similarly, *anthropectomy* refers to "the error of denying that an animal has a certain characteristically human capacity when in fact it does have that capacity" (Andrews & Monsó, 2021; see also Andrews & Huss, 2014). Both

are anthropocentric biases because they inadvertently encourage underattribution.

However, inasmuch anthropomorphism might promote false positives (i.e., ascribing a psychological property to an animal when it lacks that property, the “anthropomorphic error”), anthropodenial and anthropectomy might promote false negatives (i.e., denying a psychological state to an animal who actually has that mental state) (see Andrews & Huss, 2014 for a fine discussion of the asymmetry of both kinds of errors in favor of anthropomorphist ones). Finally, rather than gauging the degree to which MC might benefit our understanding of animal minds, Andrews, among others, proposes that it would be more useful “...ignore it as prejudice, and instead work on developing methods for testing the applicability of specific properties” (2009, p. 52). In the same sense, it has been objected that MC is redundant or superfluous: strict adherence to facts (*empiricism*) (Sober, 2005) or to evidence (*evidentialism*) (Fitzpatrick, 2008) is all that is required to choose between alternative interpretations, avoiding anthropomorphic errors. On the other hand, there is a fertile avenue for an anthropomorphic approach. de Waal (1999) refers to *animalcentric anthropomorphism* as the attempt to understand animals “on their own terms,” that is, to understand their behaviors “within the broader context of the specie’s habits and natural history”. In this sense, anthropomorphism has a positive heuristic value.

The MC has also been interpreted as equivalent to a principle of epistemological simplicity that is adopted by default: but a strictly behaviorist (non-mentalistic) explanation could have even greater complexity. On the other hand, as de Waal warns: “The word *simple* is not as simple as it seems. It means different things in relation to different species” (2016, p. 55). MC has often been interpreted as recommending adopting parsimonious ontological assumptions, but positing *fewer* entities should not be confused with positing *less complex* entities or *fewer cognitive processing costs*. These are different kinds of *cognitive parsimony* (see de Waal, 1999). But, when dealing with related species, it might be evolutionarily parsimonious to attribute similar cognitive processes to explain similar behaviors (de Waal, 2016, p. 269). This would be a kind of parsimony based on evolutionary foundations, or *evolutionary parsimony* (de Waal, 1999). Both types of parsimony, *cognitive* and *evolutionary*, try to avoid the errors typical of each type of bias: the first, anthropomorphic, the second, anthropocentric. Now then, in the case we should choose between the risks of overestimating or underestimating the psychological capacities of animals, we must take into account that new evidence can discredit an anthropomorphic error, while denying the presence of one or another

ability is not an invitation to skepticism or agnosticism, still less to continue investigating.

MC was also associated with the standard of objectivity in the context of animal studies. In this sense, it is interpreted that it promotes the search for a dispassionate, de-subjectified approach, without affection, without contact or personalized bond with the study subjects (e.g., the type of objections that Jane Goodall's methodology and her lexicon received, in the sections following). This way of understanding scientific objectivity as a "critical distance" as well as the methodological styles it encourages, could also explain "excessive concerns about anthropomorphism" (de Waal, 2016, p. 61). But naïve anthropomorphism should not be confused with justified psychological attribution: "It is anthropomorphic to call the lion the King of Beasts, but not to talk of him as moved, now by fear, now by curiosity, now by territorial anger" (Midgley, 1978, p. 74). On the other hand, in this context, objectivity should be understood as "simply fair, unbiased, and impartial" (Midgley, 2001, p. 753).

Finally, Buckner (2013) points out that the precautions encouraged by MC must be symmetrical, that is, applied equally to humans. As we know, not all human actions presuppose sophisticated cognitive mechanisms. Furthermore, simple mechanisms also operate at the basis of some complex human behaviors (Shettleworth, 2010). Morgan himself warned that

[T]o interpret animal behavior one must learn also to see one's own mentality at levels of development much lower than one's top-level of reflective self-consciousness. It is not easy, and savors somewhat of paradox" (Morgan 1930, 250).

This variety of anthropocentric bias is *anthropofabulation*: the "tendency to link competency criteria for cognitive abilities to an exaggerated sense of typical human performance" (Buckner, 2013, p. 853). This bias is the result of the combination of *semantic anthropocentrism* with *confabulation*, which consists of "overestimating our cognitive sophistication" (p. 860)<sup>12</sup>. This specific semantic bias, as we will see in section 7, consists of refining psychological concepts in such a way that they refer only to "genuine" cases of each psychological capacity. This first step, usually unnoticed in many debates between optimists and skeptics about animal cognition, generates "anthropofabulous taxonomies" about the key terms

12. This bias not only influences animal cognition: applied all the time to our own auto-attributions is *confabulation*. So, *anthropofabulation* combines *semantic anthropomorphism* with *confabulation*.



of the debate. The philosophy of the human mind is plagued by such taxonomies. Their comparative effects on animal cognition should not be surprising.

As Buckner points out, Hume identified both anthropofabulation and its appropriate corrective (1739):

When any hypothesis... is advanced to explain a mental operation, which is common to men and beasts, we must apply the same hypothesis to both; and as every true hypothesis will abide this trial, so I may venture to affirm, that no false one will ever be able to endure it. The common defect of those systems, which philosophers have employed to account for the actions of the mind, is, that they suppose such a subtlety and refinement of thought, as not only exceeds the capacity of mere animals but even of children and the common people in our own species." (T1.3.16.3; SBN 177).

I will identify other expressions of semantic bias in animal cognition research in the section 7.

## 7. ANTROPOCÉNTRIC METHODOLOGIES

The bias we are examining also manifests itself in the methodologies chosen for studying animal cognition. Firstly, the contrast between so-called "romantics" and "killjoys" corresponds more or less directly to field versus laboratory research and with different base disciplines from those who carry them out (Andrews, 2020b). They are also reflected in the experimental design in comparative studies: what abilities we want to study, in what species, and through what types of strategies and tasks (cf. Boesch, 2007). It has been usual to choose skills that humans are notoriously skilled at (e.g., "the magical wells of our species, like language" (de Waal, 2016, p. 22), or that are well identified by human psychology (Buckner, 2013). But why not also study those in which animals are specialists?

The same bias is observed in the performance of arbitrary laboratory tasks, not linked to the problems that animals face in their natural environments (Waal, 2009). Another clear example is the comparative studies on the theory of mind in chimpanzees and children: while children must deal with humans, known individuals of their species, apes face members of another species, unknown humans. As de Waal points out, that way: "All we are testing is the ape's theory of the *human* mind." (2016, p. 146). For their part, dramatically different conditions could lead to incomparable conclusions. Thus, by both its design and its motivations, research in

comparative cognition often ended “turning the study of cognition into a contest” (de Waal, 2016, p. 248).

The cognitive tests should not be identical when they take into account the behavioral ecology of each species: the animals should be tested in less artificial environments and through a “co-specific approach”, vgr. anthropoids with anthropoids, or testing their abilities with the participation of a human environment with which they were familiar and comfortable, ensuring *a priori* interaction with the researcher (Vitale & Branchi, 2023). As has been noted, primates distinguish caregivers from researchers, and researchers, by their greater or lesser previous relationship with them. In this way, the degree of habituation and trust in the researcher makes a big difference in the results of the tests (Boesch, 2007). Also it must be compare species taking into account developmental trajectories (Shettleworth, 2012).

For its part, experimental evidence must be complemented by what can be obtained through observation of behaviors in natural habitats. In such scenarios, social behavior can be better studied, which is very revealing in certain species, in contrast to the case of animals that have grown in captivity. On the other hand, anecdotal reports, so insistently questioned for their limitations, offer interesting information for investigating animal cognition: they can provide an accurate description of rare or difficult-to-observe behaviors and generate new hypotheses and even novel areas of research<sup>13</sup>. Indeed, the observation and recording of a “disconcerting” or strange behavior can serve as a starting point for an investigation as long as it comes from expert observers and guide a rigorous investigation that rules out a casual phenomenon. That is, anecdotes should serve as a basis for data collection and not more anecdotes (Andrews, 2020b). These could generate hypotheses to be investigated (Bates & Byrne, 2007).

The “Goodall case” summarizes the impact of the anthropocentric standard on her investigative strategies: observational methodology, anecdotal records, the personal knowledge of the individuals studied, and “personalized” treatment of animals yielded lasting results in primatology. However, she was criticized for these same practices:

Godall’s apes had personalities, emotions, and social agendas. She did not unduly humanize them, but she related what they did in unpretentious prose that would have been perfectly normal for a day at the office

13. These authors examined publications between 2000 and 2016 in four high-impact journals in Primatology. They revealed the discredit that anecdotal reports still suffer compared to studies based on experimental designs and quantitative methods.

but was unorthodox with regard to animals. It was a huge improvement over the tendency at the time to drown behavioral descriptions in quotation marks and dense Jargon in order to avoid mentalistic implications. Even animal names and genders were often avoided. (Every individual was an “it.”) Goodall’s apes, in contrast, were social agents with names and faces. (de Waal, 2016, p. 170).

Imanishi, the founder of Japanese Primatology, also gave names to his animals. This fact precisely reflected a degree of interaction, and at the same time, special attention was paid to the personality characteristics, behaviors, and social ties of each individual studied. The “combat” of paradigms that occurred between Western and Japanese primatology in the middle of the last century that de Waal reconstructs, reflects how anthropocentric bias was affected by positive cultural factors: “The study of animal behavior in Japan has never been contaminated by feelings of superiority or an aversion to acknowledging humanlike characteristics in animals” (2003, p. 294).<sup>14</sup> This allowed him to discover many characteristics in primates that were largely unexplored and then resisted by Western primatology: for example, identifying each individual, studying the complex social relationships of primate communities, and describing the patterns of learning and cultural transmission.<sup>15</sup> As we will see below, personalized treatment and the use of proper names are part of a non-anthropocentric way of using language towards animals that reflects, in turn, a different way of trying to understand their point of view (Andrews, 2020b). Referring to similar approaches and strategies, de Waal offers a revealing first-person testimony: “I can’t count the number of times I have been called naïve, romantic, soft, unscientific, anthropomorphic, anecdotal, or just a sloppy thinker” (2016, p. 265). Such kind of disapprovals originates in the bundle of anthropocentric biases we are analyzing here.

On the other hand, it is notable that studies on animal cognition focused largely on primates, another reflection of our obsession with identifying the traits that make us (or seem to make us) unique. In that sense,

14. de Waal also highlights the existence of significant imbalances (a lot of them are varieties of “human-centrism”) that contaminates the dialogue between cultures in the respective scientific communities (e.g., the privileges and prejudices that derive from speaking the dominant language, the neglect of the contributions from other traditions once assimilated into the Western mainstream, the tendency of the intellectual canon to underestimate other research traditions, etc.).

15. de Waal (2006) makes another interesting point related to anthropocentric inheritance in studies of animal cognition: identifying cognition with individual rather than “interindividual” capacities. The latter, such as empathy, helping behaviors, compassion, reconciliation, etc., have not been seen for a long time as “a sign of intelligence.”

“chimpocentrism” should be seen as “a mere extension of anthropocentrism” (de Waal, 2016, p. 162), that is, as another manifestation of a restrictive program in studies on animal cognition.

## 8. SEMANTIC ANTHROPCENTISM

In comparative studies on cognition in different species, it is not possible to avoid the discussion about which concepts serve to guide research (Andrews, 2009; Buckner, 2013; Figdor, 2018). It is clear that in our ordinary dealings with some animal species, we spontaneously employ the concepts of folk psychology. It is a well-described phenomenon and partly captured by the literature on anthropomorphism in its folk variety or naïve anthropomorphism. Now, in comparative psychology and biology in general, even at different explanatory levels, the capacities that one wishes to investigate are also conceptualized through the usual predicates of human psychology (cf. Figdor, 2017).

The debate about the usefulness and risks of using anthropomorphic language in animal studies oscillates between the most varied extremes: justified versus biased; useful versus harmful; literally appropriate versus simply a stopgap heuristic, etc. From the perspective we are interested in, the use of psychological predicates beyond the human case divides “sensibilities”, as we saw above, between anti-exceptionalists or continuists, who defend such predicates maintain their literal meanings beyond the human mind (see Figdor, 2018) and exceptionalists who maintain that such predicates cannot be applied strictly sense beyond the human case. From this last perspective arises the anthropocentric semantics of psychological concepts, which distinguishes between literal meanings for the human case and different non-literal interpretations (deflated or “minimal”, instrumental, analogical, metaphorical, etc.) when applied to non-human animals.

It is clear that to avoid both anthropomorphism and anthropocentrism, the use of psychological concepts towards animals requires specifying or partially modifying their meaning, to apply them appropriately, that is, taking into account the particular characteristics of each species. Now, as Andrews (2009) points out, the concepts of the everyday psychological repertoire should not be interpreted as establishing necessary and sufficient conditions for their application, even when we apply them to ourselves. That is, it would not be correct to assign them a meaning that can be precisely displayed because they do not express a univocal or homogeneous concept. Rather, they are used as “general terms.” Thus, when

we talk about emotions, intentions, purposes, empathy, or friendship, the phenomenon denoted admits varieties, characteristic forms, and degrees.

On the other hand, it is curious that when considering “authorizing” its use to refer to abilities in animals, its over-intellectualized version is frequently used, as we have already seen in the previous section: that is, a belief is characterized as an epistemic attitude with propositional content, verbalizable by the being that possesses it, and inserted in a network of other beliefs articulated among themselves inferentially, that the subject knows that he possesses and of which he can give an explicit account, etc. One version of this approach is exemplified by Davidson’s arguments for whom we should not attribute a psychological state or capacity to a creature if it does not itself possess the concept that could allow it to self-attribute that state or capacity to itself or another, i.e., if said concept is not part of its conceptual framework. Now, this could be a case of attribution quite similar to that presented to the ethnographer and ethnologist. The first records the behaviors of a human group through the conceptual repertoire of that human group, that is, through its “native categories.” The ethnologist, on the other hand, does so using his categories, because he seeks to identify cultural similarities and differences, and needs a conceptual framework that makes comparison possible. Now, the speculative philosopher, when dealing with animal minds, usually applies a variety of “auto-anthropology” that is doubly naïve, because not only does he not take into account the scientific knowledge available about them, but only the refinement of beliefs or intuitions from our popular cognitive ethology, usually a personal selection of them. Philosophers of this style tend to “declare that various propositions are true for the sole reason that they seem tremendously obvious to *them*” (Dennett, 2013, p. 99). But there are other alternatives.

Now, as Figdor (2018) points out, the relevant psychological concepts, whatever their roots in folk psychology, are those that, lastly, establish scientific disciplines. According to his diagnosis, after overcoming a century dominated by behaviorist restrictions and the survival of exceptionalist metaphysics, a massive process of conceptual change would finally be occurring, guided by changes in scientific discoveries about the animal mind. This process should lead to the end of anthropocentrism, both psychological and semantic. Conceptual change requires specifying and diversifying psychological terms, but it could also consist of associating pre-existing terminology with entirely new meanings, allowing that “... emerge from patterns observed in the data themselves. In short: first the data, then the words.” (Döring & Chittka, 2011, p. 94). That is, “If findings do not fit under the umbrella of a particular and established cognitive

construct they may still be close by and deserve a novel term. This is certainly better than repeatedly (and more or less forcefully) adjusting the size of the umbrella.” (Döring & Chittka, 2011, p. 93).

Semantic anthropocentrism also expresses varying degrees of skepticism about animal minds. One of its manifestations is the recourse to “as if” rhetoric that accompanies the use of mental predicates on animals. With the same purpose, some characteristic epithets, which are the fruit of successive redefinitions of pre-existing notions, are also used. “Genuine”, truly, or “plain”, i.e., “genuine imitation”, or “plain theory-of-mind” (de Waal, 2016, p. 126) distinguishes of “quasi” or “proto” or “not-plain” abilities. Another common strategy is to apply the same psychological terms to animals but in quotation marks, whether explicit or implicit, or using asterisks (to distinguish, for example, our “genuine beliefs” from animal beliefs\*, which share a common element, for example, “reliable differential response dispositions,” but differ in other critical elements (see Figdor, 2018). In these cases, only the human variety of the cognitive phenomenon at issue it assumes to be genuine. In contrast, one can admit a shared concept with differences relative to each species without assuming a canonical definition as a standard or criterion. That is, without adopting the “imperialist pretensions” of exceptionalism (Figdor, 2018, p. 81).

According to de Waal, all these resources reflect an “obsession with semantics, definitions, and redefinitions, and... the movement of goalposts” (de Waal, 2016, p. 157). Although seemingly motivated by a search for greater precision in the meaning of the terms, sometimes they only cover the desire to differentiate the human version from any other. Philosophical literature is rife with these redefinition exercises to distinguish a given capacity as long as it is realized voluntarily, with greater sophistication, flexibility, insight, purpose, and/or capacity for refinement, from that same capacity when it lacks some or all of those properties. In terms of Buckner: “...reliance on such honorifics (genuine, real, full-bawn, etc.) has a way of leading to constantly shifting goalposts; every time an animal or artificial system satisfies a previously specified benchmark, the critic can simply endorse a yet more restrictive interpretation of ‘real’ or ‘genuine’ and push the borderline ever-closer to the uppermost limits of human performance—and possibly even beyond.” (2023).

An even more skeptical conception of animal cognition is reflected in the tendency to avoid some words from the psychological lexicon, introducing a frankly depsychologized jargon in their replacement. As Andrews states: “Instead of describing two individuals as “friends”, many prefer to speak of their “affiliative relationship”. And rather than using emotional

state descriptions such as “happy” or “sad” or “depressed” to describe an animal, jargon may be introduced.” (2009, p. 51). When it comes to chimpanzees, instead of characterizing their behavior similar to ours as a “kiss,” “mouth-to-mouth contact” is recommended (de Waal, 2016, p. 26). This caution occurs especially with psychological terms because, it is argued, they could carry anthropomorphizing attributions. To that extent, the attribution to animals of fear, anger, stress, and the like can be admitted, but not thoughts, intentions, purposes, moods, or personality traits. This type of strategy has a negative effect to the extent that it imposes “unjustified linguistic barriers that fragment the unity with which nature presents us” (de Waal, 2016, p. 26). As a result, the predictive power of the statements that can be made using these terms is limited.

In scientific publications, the recommendation to refer to animals using non-agential neutral pronouns, available in English, when the sex of the animal is known, reflects a treatment of them as if they were objects or things (see Andrews, 2020b). A frequent manifestation of this linguistic speciesism shows in the use of possessive pronouns to assimilate our relationship with animals to an object bonds, more specifically, to an owner: e.g., my dog. These linguistic practices, rather than ensuring a standard of objectivity, reflect “a robotic view of animals” (de Waal, 2016).

## 9. CONCLUSIONS

de Waal calls “neo-creationism” the conception that half-accepts evolution in the sense that “evolution stops in the head.” There would be a “magical” leap, which means that between our mind and other types of minds there is a “spreader, gap or abyss” (2016, p. 122). According to de Waal, although the predominant assumption in the natural sciences is continuity, the discontinuist view still prevails in the social sciences and humanities. I agree with this diagnosis.

When de Waal (2016) asks: “Are we smart enough to know how smart animals are?” suggests that our highly overestimated intelligence went through extended periods of self-affirmation that had as a direct correlate a severe inability (not just intellectual) to understand animal cognition. This inability, to the extent that it derives from a persistent bias, the *anthropocentric bias*, continues to influence our approach to the topic. Although biases are part of science, the best way to minimize their impact is to identify how they influence our understanding (Andrews, 2020b). Devoting reflective efforts to anthropocentrism, a supposedly vestigial concept, reveals its interest when it notes its different dimensions and

metamorphoses in the humanistic and natural disciplines. That task is not yet completed in all its refinement. It is worth keeping in mind that, until very recently, the attribution of mental abilities to animals was illegitimate. As de Waal (2016) states, it is mainly due to Cognitive Ethology founded in the 80's of the last century that the concept of animal cognition stopped being interpreted as a contradiction in terms.

Critical reflection on the history and evolution of studies on animal cognition contributes to understanding the effects of this bias. That is the enormous value of books like de Waal's (2016). A direct remedy against them will come from the convergence between the perspectives, methods, and languages of the different disciplines that address the phenomenon, as proposed by Andrews (2020b). The imprint of the anthropocentric canon over all disciplinary studies on animal cognition gives additional foundations to the cooperative effort between natural, social, and human sciences. That effort, relatively incipient, has begun to bear fruit. This review attempts to contribute to a critical reflection that both facilitates and stimulates that interdisciplinary confluence.

Meanwhile, a phenomenon as marvelous to human research curiosity as the diversity of species and the diversity of their capacities, for different causes and in different ways, risks becoming more and more a bookish concept rather than a real phenomenon. It would be a sign of intelligence on our part to dispel prejudices that allow us to understand them better. Are we smart enough to do it? Are we smart enough to stop being anthropocentric?

## REFERENCES

- Allen, C. (2017). On (not) defining cognition. *Synthese*, 194, 4233-4249.
- Andrews, K. (2009). Politics or metaphysics? On attributing psychological properties to animals. *Biology and Philosophy*, 24, 51-63.
- Andrews, K. (2020a). *The Animal Mind. An Introduction of Animal Cognition*. London & New York: Routledge.
- Andrews, K. (2020b). *How to Study Animal Minds*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108616522>
- Andrews K., Huss B. (2014). Anthropomorphism, anthropectomy, and the null hypothesis. *Biol. Philos.*, 29(5), 711-729.
- Andrews, K., Susana, M. (2021). Animal Cognition. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (Ed.). <https://plato.stanford.edu/archives/spr2021/entries/cognition-animal/>



- Bastian, B., Loughnan, S., Haslam, N., Radke, H. R. M. (2012). Don't Mind Meat? The Denial of Mind Used to Human Consumption. *Personality and Social Psychology Bulletin*, 38(2), 247-256.
- Bates, L. A., Byrne, R. W. (2007). Creative or Created: Using Anecdotes to Investigate *Animal Cognition*. *Methods*, 42, 12-21.
- Boesch, C. (2007). What Makes us Human (Homo sapiens)? The Challenge of Cognitive Cross-Species Comparison. *Journal of Comparative Psychology*, 121, 227-240.
- Borchert, R., Dewey, A. R. (2023). In praise of animals. *Biol. Philos.*, 38, 24. <https://doi.org/10.1007/s10539-023-09912-2>
- Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.
- Bräuer, J., Hanus, D., Pika, S., Gray, R., Uomini, N. (2020). Old and New Approaches to Animal Cognition: There Is Not 'One Cognition'. *Journal of Intelligence*, 8(3), 28.
- Buckner C. (2013). Morgan's Canon, meet Hume's Dictum: avoiding anthropofabulation in cross-species comparisons. *Biol Philos*, 28(5), 853-871.
- Buckner, C. (2017). Understanding Associative and Cognitive Explanations in Comparative Psychology. In K. Andrews and J. Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* (pp. 409-418). London: Routledge.
- Buckner, C. (2023). Black Boxes or Unflattering Mirrors? Comparative Bias in the Science of Machine Behavior. *The British Journal for the Philosophy of Science*, 74(3), 681-712.
- Chittka, L, Rossiter, S. J., Skorupski, P., Fernando, Ch. (2015). What is comparable in comparative cognition? *Phil. Trans. of the Royal Society B*, 36, 2677-2685.
- Coetzee, J. M. (1999). *The Lives of Animals*. Princeton: Princeton University Press.
- Darwin, C. (1859). *On the Origin of Species*. London: John Murray.
- Darwin, C. (1854/1989). "Letter to Joseph Hooker", Jun 27, 1854 (1573). *The Correspondence of Charles Darwin, vol. 5 1851-1855*. Cambridge: Cambridge University Press. Darwin Correspondence Project.
- Darwin, C. 1982 [1871]. *The Descent of Man, and Selection in Relation to Sex*. Princeton, NJ: Princeton University Press.
- Davidson, D. (1985). Rational Animals. In E. LePore and B. McLaughlin (Eds.), *Actions and Events*. Oxford: Blackwell.
- Dennett, D. (1998). *The Intentional Stance*. Cambridge, Massachusetts: MIT Press.

- Dennett, D. (2013). Kinds of things - Towards a Bestiary of the Manifest Image. In D. Ross, J. Ladyman, & H. Kincaid (Eds.), *Scientific Metaphysics* (96-107). Oxford: OUP.
- Descartes, R. (1991). *The Philosophical Writings of Descartes: Volume III, The Correspondence*. Cambridge: Cambridge University Press.
- de Waal, F. B. (1999). Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. *Philosophical Topics*, 27, 255-280.
- de Waal, F. B. (2003). Silent Invasion: Imanishi's Primatology and cultural biases in science. *Anim. Cogn.*, 6, 293-299. <https://doi.org/10.1007/s10071-003-0197-4>
- de Waal, F. B. (2006). *Primates and Philosophers*. Princeton, NY: Princeton University Press.
- de Waal, F. B. (2009). Darwin's last lough. *Nature*, 460, 175. <https://doi.org/10.1038/460175a>
- de Waal, F. B. (2016). *Are we smart enough to know how smart Animals are?* New York: W. W. Norton & Company.
- de Waal, F. B., Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, 14(5), 201-207.
- Döring, T. F., Chittka, L. (2011). How Human are Insects, and does it matter? *Formosan Entomol*, 31, 85-99.
- Figdor, C. (2017). On the proper domain of psychological predicates. *Synthese*, 194(11), 4289-4310.
- Figdor, C. (2018). *Pieces of mind: The proper domain of psychological predicates*. Oxford and New York: Oxford University Press.
- Figdor, C. (2021). The psychological speciecism of humanism. *Philosophical Studies*, 178, 1545-1569.
- Fitzpatrick, S. (2008). Doing away with Morgan's Canon. *Mind & Language*, 23, 224-246.
- Fitzpatrick, S. (2018). Against Morgan's Canon. In K. Andrews and J. Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds*, (pp. 437-447). New York: Routledge.
- Flack, J. C., de Waal, F. B. (2000). 'Any animal whatever': Darwinian Building Blocks of Morality in Monkeys and Apes. *Journal of Consciousness Studies*, 7(1-2), 1-29.
- Glock, H. J. (2012). The anthropological difference: What can philosophers do to identify the differences between human and non-human animals? *Royal Institute of Philosophy Supplement*, 70, 105-131.
- Heil, J. (1982). Speechless Brutes. *Philosophy and Phenomenological Research*, 42, 400-406.

- Holroyd, J., Sweetman, J. (2016). The heterogeneity of Implicit Bias. In Michael Brownstein & Jennifer Saul (Eds.), *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology* (pp. 80-103). New York, USA: Oxford University Press.
- Horta, O., Albersmeier F. (2020). Defining speciesism. *Philos Compass*, 15(11), 1-9.
- Hume, D. (1739/2000). *A Treatise on Human Nature*. Oxford: Oxford University Press.
- Kahneman, D., Slovic, P., Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Lovejoy, A. O. (1936). *The great chain of being: a study of the history of an idea*. Cambridge, MA: Harvard University Press.
- Malcolm, N. (1972-73). Thoughtless Brutes. *Proceedings and Addresses of The American Philosophical Association*, 46, 5-20.
- Mayr, E. (1988). *Toward a new philosophy of biology: Observations of an evolutionist*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- McDowell, J. (1996). *Mind and World*. Cambridge, MA: Harvard University Press.
- McDowell, J. (2009). *Having the World in View. Essays on Kant, Hegel, and Sellars*. Cambridge, Mass: Harvard University Press.
- Midgley, M. (1978). *Beast and Man*. Ithaca, New York: Cornell University Press.
- Midgley, M. (2001). Being objective. *Nature*, 410(6830), 753.
- Morgan, C. (1903). *An introduction to comparative psychology*. London: Walter Scott Pub. Co.
- Morgan, C. (1930). Autobiography of C. Lloyd Morgan. In Murchison C (Ed.) *History of psychology in autobiography* (Vol. 2, pp. 237-264). Worcester: Clark University Press.
- Noske, B. (1993). Great Apes as Anthropological Subjects -Deconstructing Anthropocentrism. In Cavallieri, P. and Singer, P. (Eds.) *The Great Ape Project. Equality beyond Humanity* (pp. 258-268). New York: St. Martin Press.
- Osborne, C. (2007). *Dumb Beasts and Dead Philosophers. Humanity and the Human in Ancient Philosophy and Literature*. Oxford: Clarendon Press.
- Povinelli, D. J. (2004). Behind the ape's appearance: escaping anthropocentrism in the study of other minds. *Daedalus*, 133(1), 29-41
- Preece, R. (2005). *Brute souls, happy beasts, and evolution: the historical status of animals*. Vancouver, Canada: UBC Press.
- Rigato, E., Minelli, A. (2013). The great chain of being is still here. *Evolution: Education and Outreach*, 6, 18.

- Ritvo, H. (2007). On the Animal Turn. *Daedalus*, 136, 118-122.
- Romanes, G. (1882). *Animal Intelligence*. London: Kegan Paul.
- Shettleworth, S. J. (2010). Clever animals and killjoy explanations in comparative psychology. *Trends in Cognitive Science*, 14, 477-48.
- Shettleworth, S. J. (2012). *Fundamentals of Comparative Cognition*. Oxford: Oxford University Press.
- Singer, P. (2009[1975]). *Animal liberation*. New York: Harper Collins.
- Sober, E. (2005). Comparative psychology meets evolutionary biology. In L. Daston & G. Mitman (Eds.), *Thinking with animals: new perspectives on anthropomorphism* (pp. 85-99). New York: Columbia University Press.
- Suárez-Ruiz, J. E. (2021). El antropocentrismo como norma filosófica implícita: una revisión a la luz de las condiciones de posibilidad de la Covid-19. *Anacronismo e Irrupción*, 12(22), 78-96.
- Ullrich, R., Mittelbach, M., Liebal, K. (2017). Scala naturae: The impact of historical values on current 'evolution of language' discourse. *Journal of Language Evolution*, 3(1), 1-12.
- Vasilieva, O. (2019). Beyond "uniqueness": Habitual traits in the context of cognitive-communicative continuity. *Theoria et Historia Scientiarum*, 16, 129-150.
- Vitale, A. & Branchi, I. (2023). 'Are my animals smarter than yours?': Concerns in comparing animal intelligence. *Acta Philosophica*, 1(32), 75-94.
- Wittgenstein, L. (2009). *Philosophical Investigations*. Oxford: Wiley Blackwell.
- Wynne, C. (2004). The Perils of Anthropomorphism. *Nature*, 428, 606.

eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.31231>

## EVOLVED AND CULTURAL INTUITIONS. HIGHLY SPECULATIVE REMARKS ON THE ORIGINS OF OUR SENSE OF FAIRNESS

### *Intuiciones evolutivas y culturales. Comentarios fuertemente especulativos sobre los orígenes de nuestro sentido de justicia*

Rodrigo BRAICOVICH   
Universidad Nacional de Rosario, CONICET  
rbraicovich@gmail.com

Recibido: 6/03/2023    Revisado: 17/04/2023    Aceptado: 23/06/2023

**ABSTRACT:** The question of whether the sense of fairness constitutes an exclusively human trait has been answered mostly from two polar positions: the first one unambiguously affirms such exclusivity, thus denying the relevance of cognitive ethology to understand our evaluations of justice; the second one, on the contrary, postulates the existence of a (proto) sense of fairness in non-human animals, strongly related to ours, which would make cognitive ethology highly relevant to understand the mechanisms on which our evaluative practices are based. From a position of extreme caution in relation to the possibility of (eventually) offering concrete evidence in favor of innatist theses such as the one I will defend here, I will suggest that i) in line with the rupturist positions, it is possible to preserve the human exclusivity of the sense of justice, ii) in line with the continuist positions, the relevance of studies coming from cognitive ethology is guaranteed, insofar as (*ex hypotesi*) our evaluative practices often

take as input innate psychological dispositions shared with other species. Finally, I will suggest that the concept of rationalization is central to determine in each case the possible articulation between innate dispositions and explicit justifications.

*Keywords:* fairness evaluations, moral psychology, phylogenesis, justice, rationalization.

**RESUMEN:** La pregunta acerca de si el sentido de justicia constituye un rasgo exclusivamente humano ha sido respondida mayormente desde dos posiciones polares: la primera de ellas afirma sin ambages dicha exclusividad, negando con ello la relevancia de la etología cognitiva para comprender nuestras evaluaciones de justicia; la segunda, por el contrario, postula la existencia de un (proto) sentido de justicia en animales no humanos, fuertemente emparentado con el nuestro, lo cual volvería a la etología cognitiva sumamente relevante para comprender los mecanismos sobre los que se basan nuestras prácticas evaluativas. Desde una postura de extrema cautela en relación con la posibilidad de ofrecer (eventualmente) evidencia concreta en favor de tesis innatistas como la que defenderé aquí, sugeriré que i) en línea con las posturas rupturistas, es posible preservar la exclusividad humana del sentido de justicia, atendiendo al hecho de que nuestras prácticas evaluativas presuponen el uso de conceptos, y que ii) en línea con las posturas continuistas, la relevancia de los estudios provenientes de la etología cognitiva se halla garantizada, en la medida en que (ex hypotesi), nuestras prácticas evaluativas frecuentemente toma como *input* disposiciones psicológicas innatas compartidas con otras especies. Sugeriré, por último, que el concepto de racionalización es clave para determinar en cada caso la posible articulación entre disposiciones innatas y justificaciones explícitas.

*Palabras clave:* evaluaciones de justicia, psicología moral, filogénesis, justicia, racionalización.

## 1. INTRODUCTION

The idea that we are born with certain dispositions that are shared by every human being, or that there is, in other words, such a thing as a human nature, has been debated in Western thought since at least the Hellenistic period. The idea that there are patterns of behavior that derive from an innate makeup has been resisted fiercely on account both of its methodological flaws, its political underpinnings and its ethical consequences (see, for instance, Dupré, 2001; Gould, 1996; Levins & Lewontin, 2009; Sahlin, 2008), but the very denial of human nature has, in turn, been

met with a similar reaction, since it has been contested not on methodological grounds but mainly because of the alleged political agenda that has driven such a denial (see Haidt, 2012; Mosterín, 2011; Pinker, 2002; Wilson, 1997). Contrary to what the news portals and the media might suggest with their permanent insistence on every alleged discovery of human traits that might be present in other species or that have somehow been proven to be universal, the heyday of the investigations concerning human nature is long over. After the first wave of sociobiology had passed, the apparently accelerated and exponential developments in genetics, neurology, developmental psychology, cognitive ethology and biological anthropology revived the promise that, led by the hand of evolutionary psychology, a new and definitive comprehension of the evolutionary origins of human nature would be possible. Things happened in the way, unfortunately: both the modularity thesis and the basic emotions theory (two not so often acknowledged pillars of the models of human nature en vogue by the end of the 20<sup>th</sup> century) came under heavy fire, paving the way for the resurgence of cultural and constructionist approaches (see Brooks et al., 2017; Buller & Hardcastle, 2000; Feldman Barrett, 2017; Feldman Barrett & Russell, 2015; Fox & Friston, 2012; George & Sunny, 2019; Palecek, 2017; Prinz, 2006; Suhler & Churchland, 2011). The quest for human nature seems thus to be facing its most serious crisis since Europe discovered the American and African Other, and the prospect of understanding the evolutionary basis on which some of the (allegedly) most recurrent traits in humans seems to dwindle as the methodological objections to such an enterprise keep piling up.

Does this mean that we should abandon all hope of ever settling the question concerning whether there is a substantive and shared ground that explains at least part of what constitutes us as humans? Are we forced to leave behind the idea that there is a part of us that is shared by non human animals and that by understanding them we may better understand ourselves – at least in what matters from a psychological, ethical and social perspective? Does ethology, in other words, have nothing to offer to philosophical anthropology? I believe that it is not necessary to arrive to such conclusions. But I believe as firmly that any conclusions we may wish to arrive at concerning the question of the existence of human nature, its evolutionary origins and the possible contributions of ethology to philosophical anthropology must remain extremely cautious in its formulation. All expressions must be preceded by “might” and nuanced by “perhaps”. And if we cannot rule out the possibility that a day may come when we may perhaps be forced to abandon such hopes, we can at least continue to explore the most plausible pathways that are still open

to pursue. One of them is the one that was opened by Frans de Waal's pioneering works on the existence of a (proto) sense of fairness in non human animals, and the aim of the following pages will be to ask whether it is a path that we can still fruitfully explore.

## 2. SENSE OF FAIRNESS: HUMAN AND NON HUMAN

The idea that our sense of fairness precedes our species and can be traced to other non human animals with which we share a common evolutionary past is not new: since Charles Darwin suggested in *The Descent of Man* (1871/1880) the presence of (what we call) moral emotions in several species other than ours, the evidence in favor of the hypothesis of a (proto) sense of justice in other animals seemed to accumulate considerably. De Waal and Sarah Brosnan, on the one hand, and Mark Bekoff and Jessica Pierce, on the other, were among the most prominent defenders of that idea. Since as far as 1991, De Waal has defended (with variations) the idea that "the human sense of justice evolved from inclinations observable in our simian relatives" (1991, p. 335). In his view, such a sense of fairness (or justice<sup>1</sup>) entails not only the capacity of being aware of social rules but also of developing expectations about how certain interactions *should* take place. De Waal is confident that at least some non human animals possess both:

I will describe behavior in chimpanzees and other primates that seems to reflect a sense of social regularity, that is, a sense of how others should or should not behave. This sense, which may be a precursor of the sense of justice, is defined here as a set of expectations about the way in which oneself (or others) should be treated and how resources should be divided, a deviation from which expectations to one's (or the other's) disadvantage evokes a negative reaction, most commonly present in subordinate individuals and punishment in dominant individuals. (1991, p. 336)

Bekoff and Pierce's more recent approach suggest a similar understanding of the phylogenetic roots of our own sense of fairness:

A sense of justice is a continuous and evolved trait. And, as such, it has roots or correlates in closely related species or in species with similar

1. For merely practical reasons, in what follows I will refer to the issues related to 'justice' as issues of 'fairness'. When I speak of a 'sense of fairness', for example, I take it interchangeable with the expression 'sense of justice' used by De Waal, Bekoff, John Rawls and others.



patterns of social organization. It is likely, of course, that a sense of justice is going to be species-specific and may vary depending on the unique and defining social characteristics of a given group of animals; evolutionary continuity does not equate to sameness. (2009, p. 115)

The implications involved in both approaches are important and far ranging: if it is truly the case that our sense of fairness is partly built on or rooted in certain dispositions or psychological mechanisms that we share with other species, then we cannot begin to understand the general structure and contents of our fairness evaluations until we have understood those shared elements. The presuppositions of this approach, however, are equally important and challenging, and it is those presuppositions that we need to delve into before we consider the possible connections between our fairness evaluations and other non human animals.

### **3. SENSE OF FAIRNESS: UNIVERSAL, INNATE, ADAPTED, HYPOTHETICAL**

What do we mean exactly when we speak of the human *sense of fairness*? Do we mean that, as Nicolas Baumard (2016) suggests, we are endowed with a spontaneous tendency, akin to our senses of smell or touch, to distinguish between fair and unfair situations? Is it an 'innate instinct', as James Wilson (1997) suggests? A general definition that might perhaps be accepted by many of those who have written on the subject would be the following: the human sense of fairness is the spontaneous and unconscious tendency to evaluate situations in terms of fairness or unfairness. As is evident, however, since a mere definition is no proof of the actual existence of the thing defined, we can ask the following: do we really have such a tendency? Is there solid evidence to support the claim of its existence? If so, is it innate? If it is indeed innate and (therefore universal), is it modular in its architecture? If so, what type of modularity do we assume it possesses? More importantly: are we in conditions to address these questions given the present state of research concerning the genetic, physiological, and neurological bases of our fairness evaluations? Or are we still doomed to conform ourselves with one more just so story, built on evolutionary game theory approaches and/or theoretical considerations? I believe this last to be the case, since the important objections that have been raised concerning both the modularity thesis and the faculty-based approaches to neurology has led us to the brink of a Kuhnian crisis from which no new paradigm has yet emerged that can serve as a general framework from within which to articulate the investigations concerning the neurological bases of our sense of justice.

That being so, perhaps the best argument we can present for the moment for the existence of an innate sense of justice is, by way of a *reductio ad absurdum*, to challenge the reader to imagine a human being not endowed with the tendency to evaluate situations in terms of fairness: would it be plausible to assume that a human being could survive without any tendency whatsoever to defend himself against any unfair treatment or any tendency to resist being pushed around or taken advantage of? Barring the alternative of such a condition being a spandrel, the by-product of another trait that is itself adaptive<sup>2</sup>, could natural selection favor the survival of such an individual? We can surely allow for such a possibility within the pages of a novel by Herman Melville, or within the tales of Jorge Luis Borges. We can even envisage it as one of the psychiatric study cases in Oliver Sacks' repertoire of singular characters, or as the (probably fleeting) result of a long and bizarre experiment in ascetic self-fashioning. But to believe that natural selection could leave the tendency to enforce fairness entirely in the hands of culture seems inconceivable in a cooperative and extremely social species as ours, since cooperation cannot evolve without the existence of the tendency to obey and enforce rules<sup>3</sup>. I believe that, at present, we are not able to produce a better account of the sense of fairness than that. There really isn't any evidence (as far as I am aware) to back up the claim that certain neural networks are responsible for the modular functioning of our tendency to evaluate situations in terms of fairness, and neither is there solid anthropological and historical evidence to support the existence of a universal sense of fairness. What we are left with is, in sum, the mere possibility of exploring theoretically an abstract proposal that may, perhaps, one day be definitely disproven or provisionally confirmed.

2. It is certainly impossible to rule out the emergence of our sense of fairness as a by-product of the emergence of consciousness, and that explanation would render the investigations in ethology mostly useless in our understanding of our fairness evaluations. What I propose to explore, however, is the (at least for now equally valid) opposite alternative.

3. It could be pointed out that in this particular instance, I am taking natural selection as operating on the individual level, but that it can also operate on a population level: alarm calls in non human primates, for example, entail a risk for the individual who utters them, but allows the rest of the population to escape a certain danger. The objection, however, would be irrelevant in this case on account of a structural difference between both scenarios: in the alarm call scenario, it is the *side effect of the possession of a certain trait* that entails a danger to its possessor; in the scenario I am discussing, the danger would come from *the absence of a certain trait* (the sense of fairness), and natural selection, as far as I know, does not select for absences.

#### 4. SENSE OF FAIRNESS, FAIRNESS CRITERIA AND FAIRNESS EVALUATIONS

If our sense of fairness consists in the tendency to evaluate certain situations as fair or unfair, it is evident that our fairness evaluations presuppose the capacity to entertain concepts in at least two distinct but complementary instances: the evaluation itself of a certain situation as *fair or unfair*, and the criteria used in that evaluation (equality, reciprocity, property, etc.). In both cases we are attributing a certain property to a state of affairs and, in order to do so, we need the corresponding concepts. While the base evaluation can be considered universal (since there doesn't seem to be evidence of any culture that lacks a concept to denote general violations of norms), fairness criteria are partly cultural and optional: they are partly cultural in that certain societies or cultures may resort to specific criteria that others do not, and they are optional in that one may limit oneself to judge a certain situation as *unfair* without stating (or knowing) the criteria on which one has based one's judgment. In either case, the base evaluation does entail the use of the concepts of 'fair' and 'unfair', and can therefore only take place (as far as we know) in the mind of a human being: there may appear to be strong parallels or analogies between human and non human animals when it comes to enforcing norms, inflicting punishment, reciprocating favors, etc., but if fairness evaluations imply the concept of fairness, those parallels and analogies can only be superficial<sup>4</sup>. But if fairness evaluations are exclusively human, what light could the study of the behavior of non human animals possibly shed on our sense of fairness? If we adopt a thoroughly rupturist approach (as, for instance, Ayala (2010), Baumard (2016), and Korsgaard, (2006) do) the answer seems to be none: ethology has no insights to offer

4. It could be argued that although non human animals lack linguistic concepts, there are other approaches to the notion of "concept" that could be relevant in this respect, such as dispositionalists approaches (a recent review of the corresponding literature can be found in Danón (2021)). That would be not be the case here, however, since I have specifically defined our sense of fairness as "the spontaneous and unconscious tendency to evaluate situations *in terms of fairness or unfairness*", which explicitly entails the classification of a situation as "fair" or "unfair". If we were to define our sense of fairness as the spontaneous and unconscious tendency to evaluate situations *in ways that we humans could equate with our classifications of fair/unfair*, the objection would certainly be relevant. But that is not what I have done (mainly because it would involve determining which is the behavioral evidence in each case that would point to those non linguistic evaluations, an endeavor that I do not feel capable of accomplishing).

to those who are interested in studying our sense of fairness or human morality in general.

I believe that such a conclusion is not warranted if we distinguish clearly between the following elements: i) our fairness evaluations, ii) the psychological dispositions that underlie them, and iii) the criteria we (may) use to justify those evaluations<sup>5</sup>. These three elements, which are sometimes confounded in other approaches (leading to either anthropomorphism or anthropodenial) are clearly distinct in kind: our fairness evaluations are the actual and conscious evaluations of certain situations as fair or unfair, and are conceptual in kind, as are the criteria we may resort to when justifying those evaluations (“His decision was unfair because I had not done anything wrong!”). The psychological dispositions that underlie those evaluations, however, do not necessarily imply the use of concepts<sup>6</sup>.

To be sure, fairness evaluations can be the result of careful and systematic deliberation: a judge may come to the conclusion that, taking into consideration certain criteria, a particular situation was unfair and therefore someone deserves to be punished. We can also come to similar conclusions after much reflection on something that has happened to us or someone else. In such cases, fairness evaluations do entail the conscious use of fairness criteria, and the study of the behaviors of chimpanzees, bonobos or dolphins appears to be completely irrelevant to further our understanding of the mental processes involved in those evaluations, in which logic (or culture) seems to have the final word. And the same seems to be the case when we deal with fairness evaluations that appear as the result of the unconscious application of fairness criteria: if, for example, I come to the conclusion that someone has evaluated a certain situation as unfair because the idea of failed reciprocity must have come into his

5. None of these elements must be confused with the *sense of fairness* itself, which is merely the (alleged) innate tendency to produce (i) *fairness evaluations* that may (or may not) be induced by the activation of certain (ii) *psychological dispositions*, and that may (or may not) be followed by a process of rationalization that resorts to certain (iii) *fairness criteria*. Even if it could be demonstrated that there is no such thing as an evolved and innate sense of fairness as I have defined it, the distinction between the remaining three elements would still hold.

6. One need not understand those innate dispositions as rigid ‘action programmes’ (as Damasio and Carvalho (2013) do); they can rather be thought of as some the building blocks that take part (alongside others) in certain behavioral responses. The fact that those dispositions interact with other (cultural or evolved) elements would then explain why individuals of the same species can react differently to a similar scenario.

mind, even if he hasn't consciously noted it, that evaluation is still (at least putatively) conceptual in nature. Fairness criteria clearly seem to come into play during unconscious processes and determine the outcome of mental processes that lead to fairness evaluations.

If we take into account the investigations that have been conducted from within dual process theories (Evans, 2008; Evans & Stanovich, 2013; Kahneman, 2012), however, fairness evaluations that are the actual result of the application of fairness criteria seem to be only half of the story, given that those evaluations can also be the result of mental processes that are (as far as we know) essentially affective in nature, and not conceptual<sup>7</sup>. If we further take into consideration the insights of Social Intuitionism (Haidt, 2001) or the interactionist approach to reason (Mercier & Sperber, 2017), conceptual evaluations appear to account for even less of the story, since the justifications we tend to give for our fairness evaluations seem to have less to do with the truth of what has led us to those evaluations and more with speculation or even fabrication. According to a long line of research that dates back to Robert Zajonc and John Bargh (both influenced by William James' early insights), a fundamental consequence of the opacity of the inferential processes that take place in our mind is that we can only see the end result of those processes, but we cannot either witness nor reconstruct the actual process that has led to a certain conclusion. According to the Social Intuitionist Model proposed by Haidt (2001), when faced with such intuitions (i.e., conclusions that are the result of inaccessible processes<sup>8</sup>) we have a tendency not to simply admit our ignorance of why we feel in a certain way about something, but rather to produce *post hoc* rationalizations (without obviously not being aware that we are rationalizing – for if we were aware that we are doing so we would be consciously lying). Hugo Mercier & Dan Sperber (2017) further suggest that those rationalizations are based not on what we hold to be the *actual* causes of our evaluations, but rather the most *plausible* ones. A coincidence between the reasons that we think or declare that have led us to condemn a certain practice as unfair and the reasons (if any) that have actually led us to that conclusion, can only be the result of chance. Our declared reasons, in sum, are only probable guesses we

7. Although the distinction between rationality and emotions (or concepts and affects) has been questioned since at least the Hellenistic period, I believe it is still useful to pinpoint the psychological dispositions that take place in the mind of non linguistic species.

8. "Intuition is the best word to describe the dozens or hundreds of rapid, effortless moral judgments and decisions that we all make every day. Only a few of these intuitions come to us embedded in full-blown emotions" (Haidt, 2012, p. 60).

produce in order to fill the void on which our sense of fairness operates, a void that seems to be filled not only by culture, but also by evolution.

## 5. CULTURAL AND EVOLVED PSYCHOLOGICAL DISPOSITIONS

As working hypotheses, I have thus far suggested that is possible that:

- i) we are born with what we may call a 'sense of fairness', which is a tendency to evaluate (in a spontaneous and unconscious manner) certain situations as fair or unfair;
- ii) that our sense of fairness may be an adaptation;
- iii) that although it may be modular (to some degree) we do not seem to be in a position at present to provide evidence concerning its physiological or neural bases.

I have also claimed that:

- iv) our fairness evaluations can be the result either of the conscious<sup>9</sup> or unconscious application of certain criteria of fairness (which are conceptual in nature), or of certain intuitions whose inner workings elude us<sup>10</sup>.

Is there room to believe that that fuzzy or dark space where our intuitions take place is inhabited by something other than merely cultural elements? Can natural selection have favored the presence of specific

9. Although there might seem to be a conflict between the proposed definition of our sense of fairness as unconscious and the admission that it may take as input the conclusions of a conscious deliberation, there is no such contradiction: the hypothesized tendency I label as sense of fairness is limited to the unconscious tendency to evaluate..., not to the actual evaluation. By way of analogy, we can think of the sensation of hunger that can sometimes lead to eat and the act of eating a specific food. Perhaps it would not be extremely inaccurate to push the analogy and think of our sense of fairness as an 'urge'.

10. After reviewing the relevant literature in the fields of psychology and philosophy, Hodgkinson et al. (2008) suggests the following as "a broadly consensual definition of intuition": "a complex set of inter-related cognitive, affective and somatic processes, in which there is no apparent intrusion of deliberate, rational thought. Moreover, the outcome of this process (an intuition) can be difficult to articulate. The outcomes of intuition can be experienced as an holistic 'hunch' or 'gut feel', a sense of calling or overpowering certainty, and an awareness of a knowledge that is on the threshold of conscious perception" (2008, p. 4). A clear and precise description of the multiple layers involved in the concept of intuition can be found in Scotto (2022).

psychological dispositions that explain certain tendencies to process information and react to it? Drawing in both cases from the tradition left behind by evolutionary psychology and by modular approaches to psychology, Haidt and his team think that that is certainly the case: nature seems to have endowed us with tendencies to react in specific manners to a set of inputs that are related to the adaptive challenges that our ancestors (both from and prior to the *homo sapiens* lineage) had to face in the past, and that explains the alleged regularities that we find below the surface of cultural diversity in terms of social rules and practices. Under this hypothesis, however, the innate and evolved dispositions that natural selection has provided us with goes far beyond the domain of fairness: we are probably born, for example, with a disposition tailored to recognize certain visual patterns (faces) and not others as evidence of the presence of a conspecific, but that disposition does not seem to have much to do with what we think of as fairness evaluations. The question, then, becomes the following: can our fairness evaluations draw on *any* evolved disposition? Or is it the case that only some of them serve as inputs to our sense of fairness? *Prima facie*, the second alternative seems to be the most plausible one: supposing, for example, that we are born with a general disposition to experience fear when faced with what appears to be a threat to our survival or well being, it would be strange to draw on that disposition and declare a situation that has produced fear in us as unfair. If it is true that we have an innate tendency to experience negative emotions towards what threatens the survival or well being of our offspring, it would be equally odd to consider such threats as unfair. I might try to do all I can to neutralize both threats and even condemn them *on other grounds*, but to label them as unfair would probably not be the most natural option.

But if it is true that our fairness evaluations do not draw on *any* evolved disposition, what determines which of the available dispositions are taken as inputs by our sense of fairness and which not<sup>11</sup>?

11. This, it must be stressed, is a descriptive question --- not a normative one: what we are asking is not whether the criteria or intuitions (cultural or evolved) that our sense of fairness relies on constitute a coherent set when taken together; nor if some of those criteria or intuitions are compatible (and others aren't) with a given abstract conception of justice that we consider on philosophical grounds to be the best available one. What we are asking, in other words, is if there is an objective (plausible) explanation to why our sense of fairness tends to rely on a certain set of dispositions (and not others) or to resort to certain criteria (and not other) – not whether it should do so from a logical, sociological or metaphysical point of view. I believe that it is precisely the failure to distinguish between both approaches that has prevented some researchers to comprehend the relevance of studying the relation between evolved dispositions and our sense of fairness.

## 6. COOPERATION, RECIPROCITY AND FAIRNESS EVALUATIONS

“She could never forget his kindness —  
he had been really remarkably kind —  
she forgot precisely upon what occasion.  
But he had been remarkably kind”  
(Virginia Woolf, *Mrs. Dalloway*)

Can the evolutionary origins of our species shed light on the previous question? Yes and no.

On the one hand, a look at our evolutionary origins certainly seems to explain the prominence and regularity that we witness in human history concerning, for example, the dynamics of reciprocity. If we assume that the cooperative nature of our species is the result of a process by which natural selection favored those individuals who showed certain dispositions that made (sustained and systematic) cooperation possible, the most basic dispositions of that sort seem to be, on the one hand, the tendency to reciprocate favors and, on the other, the tendency to adopt a negative attitude to those individuals who do not reciprocate. This does not imply, by any means, that cooperation presupposes the capacity to conceptualize rules as such: the tendencies favored by natural selection concerning cooperation may be as simple as a merely affective state that predisposes us favorably or unfavorably towards another, without any awareness of the nature and causes of that affective state. To be sure, on top of that basic initial negative attitude towards non reciprocation more complex phenomena can be built: one can feel a desire to behave aggressively towards a non reciprocator (but without actually doing so); one can effectively behave in such a manner; one can behave in such a manner while being aware of the aggressive nature of the reaction; and, at the end of the spectrum, one can be aware not only of the aggression but also of the fact that the aggression was a response to lack of reciprocation. This makes it evident that, on the one hand, complex phenomena, such as a conscious and explicit punishment of someone who has not reciprocated a favor, can be built on top of very simple evolved dispositions or tendencies, and, on the other, that those basic tendencies cannot be mistaken for their complex and cognitively demanding counterparts – a confusion that lies at the heart of the attribution of a (proto) sense of fairness to non human animals.



Jeffrey Stevens and Marc Hauser (2004) were perhaps among the first to explicitly address the question of the cognitive requirements of the mental traits that primatologists tended to attribute to non human animals. Concerning the specific case of 'reciprocal altruism' famously studied by Robert Trivers (1971), the authors claimed that several mental capacities had to be granted in a species (other than a cheater detection mechanism) for reciprocal altruism to take place: numerical quantification, temporal discounting, delayed gratification, analysis and recall of reputation, and inhibitory control. Although the authors' conclusion that reciprocal altruism must therefore be "rare if not absent among animals" (Stevens & Hauser, 2004, p. 64) seemed to condemn the investigations concerning reciprocity in animals to the museums of anthropocentrism, a more fruitful approach had already been sketched by De Waal several years before (but had remained largely ignored): in 2000, the author had put forward the notion of "attitudinal reciprocity" as a more parsimonious conceptual alternative to that of "calculated reciprocity", given that the former "is less cognitively demanding [...], because it does not assume mental score-keeping of given and received services nor expectations about appropriate return-favors, or the punishment of cheating" (De Waal, 2000, p. 260). De Waal's suggestion was later (implicitly or explicitly) taken over by several researchers: in response to Stevens and Hauser, Brosnan *et al.*, for example, resorted to the concept of attitudinal reciprocity as a phenomenon "in which individuals' responses are based on the positive feelings generated when a partner gives a favor, not on an exact accounting of favors given and received" (2009, p. 595). A similar approach was taken by Gomes *et al.* (2009), who put forward the idea that attitudinal reciprocity might involve the release of oxytocin, and by Gabriele Schino and Filippo Aureli, who suggested that De Waal's mechanism need not be limited to immediate interactions, but could be the basis of "a system of emotionally based bookkeeping that allows the long-term tracking of reciprocal exchanges with multiple partners without causing an excessive cognitive load" (Schino & Aureli, 2009, p. 5).

The substitution of the mechanism of 'calculated reciprocity' for that of 'attitudinal reciprocity' represents a paradigmatic example of the adoption of a non anthropomorphic approach to the mental lives of non human animals, in that it not also manages to put as little cognitive burden as possible on the subjects involved, but also has the additional advantage of not demanding, on the level of proximate explanations, anything more

than an affective state, rather than planification and forethought<sup>12</sup>. This last point is crucial, since it allows us to understand how reciprocity patterns can arise without any awareness of the idea of reciprocity, or, in other terms, how it is possible to find regularities without rules<sup>13</sup>. And it also explains why the dispositions that make reciprocity patterns stable can well predate our species, making research in cognitive ethology relevant to further our understanding of our fairness evaluations.

Taking into consideration how the emergence of cooperation may have favored the emergence of certain dispositions that feed into our sense of fairness, however, has important limits in explaining which evolved dispositions our sense of fairness takes as input and which ones it doesn't. And the reason for that is that reciprocity represents only one of the several criteria that humans resort to when explicitly arguing in

12. As the long debate concerning Brosnan & De Waal's famous paper 2003 shows, "minimalistic" approaches carry more plausibility when suggesting interpretations of complex mental phenomena: in this particular case, it has been suggested that instead of assuming that Capuchin monkeys have the cognitive capacity to "measure reward in relative terms" (p. 299), it is enough to resort to the "frustration effect", which (following Abram Amsel's (1958) seminal paper on the subject) has been studied extensively in several non human species (Dantzer et al., 1980; Freidin & Mustaca, 2004; Jakovcevic et al., 2013; McPeake et al., 2021; Papini, 2003; Papini et al., 2019; Stout et al., 2003; Wilton et al., 1969). Although the pattern of responses elicited by the frustration of expectations varies across species, certain observed regularities, such as the increase in aggression or agonistic behavior, seem to account more parsimoniously for at least some of the behaviors in non human animals that might at first sight be interpreted as (proto) fairness evaluations (Brauer et al., 2006; Cheney & Seyfarth, 2007; Dubreuil et al., 2006; Fletcher, 2008; Roma et al., 2006; Silberberg et al., 2009). A similar approach can be taken when proposing interpretations concerning the proximate causes of behaviors that appear on the surface to express an intention to "punish" norm violations, as has been suggested by Raihani et al. (2012) and Riedl et al., (2012).

13. De Waal's first explicit attribution of a sense of fairness to non human animals relied on this confusion: "All animals conform to social rules. That is, their conduct toward their con-specifics is to some degree predictable." (1991, p. 337). That the behavior of an individual is predictable does not entail that he is conforming to a certain rule. Charlotte Hemelrijk, for example, had already suggested that certain regularities in reciprocation that had been interpreted as evidence of Reciprocal Altruism in chimpanzees, could be explained as a side reciprocity and interchange may arise "as a side-effect from self-reinforcing aggressive interactions, spatial structure and grooming between artificial entities that lack every motivation to reciprocate" (1997, p. 190). Concerning the transactions that follow a collective hunt by a group of chimpanzees, which had previously been approached, for example, through the "meat for sex" hypothesis, Ian Gilby proposed that such transactions could be explained by resorting to the variable of harassment: the more an individual insistently claims for a share of the booty, the more chances of success he will have (2006, p. 20). From this perspective, the decision of the chimpanzee that has led the hunt to give up part of the booty does not obey a principle of reciprocity or fairness, but can be explained by factors that have nothing to do with a sense of fairness.

favor of the fairness or unfairness of a certain outcome. As Fiske and Rai (2015) have shown, for example, our fairness evaluations tend to take as input an extraordinary range of dispositions and cultural elements, which can go from violations of authority ranking to alleged divine dispositions concerning the right of some to possess more than others.

Does this mean, once again, that our fairness evaluations can be triggered by any disposition whatsoever? One of the most recent and systematic attempts to answer this question has been Haidt' Moral Foundation Theory, which claims that the whole range of dispositions<sup>14</sup> we are endowed with can be probably classified under six moral domains (care, fairness, loyalty, authority, sanctity, liberty) that are clearly distinguished by the specific evolutionary problems they were supposed to answer: in the case of the foundation of care/harm, the problem to be solved is that of protecting and caring for our offspring, in the case of the loyalty foundation, to form cohesive coalitions, and so on (Haidt, 2012, p. 139).

Haidt's theory (which explicitly draws inspiration from Fiske's Relational Models Theory) is certainly helpful in sorting out the possible evolutionary challenges faced by our ancestors and the possible dispositions that may have contributed to solving those challenges, but it faces two important difficulties: on the one hand, the criteria that Haidt and his team have proposed in order to establish the limits of each domain are not particularly clear, since at least some of the different evolutionary challenges proposed can be linked to more than one of the dispositions analyzed by them. On the other hand, the relation proposed by their model between dispositions and foundations is, as I have argued elsewhere (Braicovich, 2021), particularly vague, since the theory still fails to answer why our sense of fairness (or any of the other domains) tends to take as input certain intuitions and not others. The failure to answer that question, I believe, is, on the one hand, due to the fact that Haidt has explicitly borrowed from Sperber's distinction between the proper and the actual domain of a module, which states that a module can come to be triggered by a set of conditions that are different from the original ones under which it was shaped and proved adaptive. (So much so, in fact, that "the actual domain of any human cognitive module is unlikely to be even approximately coextensive with its proper domain from versions of modularity" (Sperber, 1994, p. 54)).

14. Although Haidt refers to them as 'emotions', I will keep the term disposition when referring to Moral Foundation Theory in order to preserve the points of contact between my proposal and his. After all, Haidt does acknowledge that the cognitive modules he is referring to have a considerable range of expressions, from mere 'flashes', to 'affective reactions' and proper 'emotions', the last of which apply to instances when a certain foundation is "activated strongly" (Haidt, 2012, p. 140).

Both Haidt's distinction between the original and the current triggers of a domain and Sperber's distinction between the proper and the actual domains of a module, in sum, are extremely unclear concerning the relation between the original/proper and the current/actual domains, which might lead one to believe that almost anything can come to trigger, for example, our sense of fairness. But that is precisely what seems to be the case when we look at the problem from a historical and ethnographical perspective, and it is that essential indefiniteness what makes both approaches relevant and interesting. And that shows, incidentally, why a look at the evolutionary history of our species (particularly concerning the emergence of cooperation) can only shed a partial light on the mental processes that underlie our fairness evaluations, leaving the rest of the work to sociology, ethnography and cultural studies<sup>15</sup>.

## 7. FINAL REMARKS

On the question of whether what we term 'sense of fairness' is an exclusively human feature or whether it can be traced to other non human species with which we share a common ancestry, I suggested the following (highly speculative) argument, which is composed of two main independent premises (1 and 4):

1. Random genetic mutations in certain individuals of non human species determined the presence in those individuals of certain innate dispositions that proved favorable to cooperative strategies.
2. As descendants of those species, we inherited those innate dispositions because of the process of natural selection.
3. Those evolved (innate) dispositions tend to trigger (in response to different stimuli) variations in our affective state (either positive or negative) which we are not (necessarily) aware of.
4. (Perhaps) for evolutionary reasons, we, as a species, are born with a tendency to evaluate intersubjective scenarios (mainly but not exclusively) in terms of 'fairness' and 'unfairness'.

15. While I doubt whether ethological studies can shed light on our sense of fairness (given that it entails the use of concepts and can be therefore deemed as an exclusively human trait – if it exists at all), I believe that cognitive ethology can certainly help us understand the evolved psychological dispositions that we have inherited from our ancestors and that still operate as (unconscious) inputs of many of our fairness evaluations. To that extent, although the contribution that cognitive ethology can make to our understanding of our fairness evaluations is only partial, it is a definite and precise contribution: we cannot possibly begin to understand those evaluations until we admit that they are too often built (among other things) on top of psychological dispositions that are shared with non human animals.

5. Those fairness evaluations can be the result of (at least) two types of mental processes: conscious deliberation and intuitions, the latter being of two types: cultural or evolved.
6. Our sense of fairness, in other words, can take as input either cultural elements or the variations in our affective state that are the result of a set of innate dispositions that were selected for due to their adaptive value concerning cooperation.

As is (controversially) evident, the main premises of this argument rest on the assumption that some version of innatism is still defensible, and that natural selection operated (at least in our evolutionary lineage) at the level of individual dispositions, rather than favoring general-purpose cognitive mechanisms, which applies not only to our 'sense of fairness', but also to the set of evolved dispositions that lie at the basis of, for example, attitudinal reciprocity or the frustration effect.

Although those dispositions can become resignified and integrated into fairness evaluation, they do not imply in themselves any notion of fairness. It certainly seems apt to think of them as the "building blocks" of the tower of morality, as De Waal has suggested, but bearing in mind that the morality (or fairness) part of the divided line starts with concepts. As the examples of attitudinal reciprocity and the frustration effect show, it is possible that the result of those dispositions is merely a rather undefined affective state with positive or negative valence and with varying intensity – affective states that do not imply either awareness or the use of concepts, but can nevertheless easily merge with other co-occurring processes that are themselves conceptual in nature. From this perspective, the effect of the frustration of our expectations, or the non reciprocation of a favor, to cite only two examples, can combine with a wide array of cultural elements (such as the multiple fairness criteria studied by Alan Page Fiske and Taje Shakti Rai (2015)) to deliver fairness evaluations that may seem at first sight to have little or nothing to do with non conceptual processes. And it is the process of rationalization that serves to obscure the indelible stamp of the lowly origins of our most sophisticated and elaborate fairness evaluations<sup>16</sup>: as Robert Solomon puts it, "much of what goes under the name of 'justice' is, one could argue, an elaboration of certain familiar feelings not dissimilar to vengeance (other, obviously, than a polar shift in 'valence') rather than a grand scheme or blueprint for the rational organization of society" (1995, p. 255).

16. We can advance a further question here: does rationalization possess an adaptive value? Mercier, Sperber and Baumard certainly seem to think so. See Baumard (2016), Mercier & Sperber (2017), and Sperber & Baumard (2012).

Understanding the central role that rationalization plays when giving account of the rationale behind our fairness evaluations, allows us to sever the causal tie between what we declare to be the motives behind those evaluations and its actual causes. Although the explicit justification of our evaluations may (by mere chance) coincide with their actual processes, we are in no way of actually knowing if they are or not, neither as philosophers, neuroscientists or psychologists, nor as regular human beings, since folk psychology, introspection and intuitive metaphysics do not seem to be better suited to explain the mental processes that our theoretical models have failed so far to account for.

But the main advantage of integrating (cultural and evolved) intuitions and rationalization into our understanding of fairness evaluations is that it allows us to approach them from a perspective that grants a role as important to culture as to evolution, preserving the exclusivity of our sense of fairness while, at the same time, acknowledging that it can frequently operate on top of evolved mechanisms that predate our species.

## REFERENCES

- Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, 55(2), 102-119. <https://doi.org/10.1037/h0043125>
- Ayala, F. J. (2010). The difference of being human: Morality. *Proc. Natl. Acad. Sci. USA*, 107(Supl. 2), 9015-9022. <https://doi.org/10.1073/pnas.0914616107>
- Baumard, N. (2016). *The Origins of Fairness: How Evolution Explains our Moral Nature*. New York: Oxford University Press.
- Bekoff, M., & Pierce, J. (2009). *Wild justice: The moral lives of animals*. Chicago: The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226041667.001.0001>
- Braicovich, R.S. (2021). La Antropología Filosófica frente al factum de la evolución. In R. López-Orellana, J. & J. Suárez-Ruiz (Eds.), *Filosofía postdarwiniana. Enfoques actuales sobre la intersección entre análisis epistemológico y naturalismo filosófico* (pp. 337-355). College Publications.
- Brauer, J., Call, J., & Tomasello, M. (2006). Are apes really inequity averse? *Proceedings of the Royal Society B*, 273, 3123-3128. <https://doi.org/10.1098/rspb.2006.3693>
- Brooks, J. A., Shablack, H., Gendron, M., Satpute, A. B., Parrish, M. H., & Lindquist, K. A. (2017). The role of language in the experience and perception of emotion: A neuroimaging meta-analysis. *Social Cognitive*

- and Affective Neuroscience*, 12(2), 169-183. <https://doi.org/10.1093/scan/nsw121>
- Brosnan, S. F., & De Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, 425, 297-299. <https://doi.org/10.1038/nature01963>
- Brosnan, S. F., Silk, J. B., Henrich, J., Marenco, M. C., Lambeth, S. P., & Schapiro, S. J. (2009). Chimpanzees (*Pan troglodytes*) do not develop contingent reciprocity in an experimental task. *Animal Cognition*, 12(4), 587-597. <https://doi.org/10.1007/s10071-009-0218-z>
- Buller, D. J., & Hardcastle, V. G. (2000). Evolutionary Psychology, meet Developmental Neurobiology: Against promiscuous modularity. *Brain and Mind*, 1, 307-325. <https://doi.org/10.1023/A:1011573226794>
- Cheney, D. L., & Seyfarth, R. M. (2007). *Baboon metaphysics: The evolution of a social mind*. Chicago: University of Chicago Press. <https://doi.org/10.7208/chicago/9780226102429.001.0001>
- Damasio, A., & Carvalho, G. B. (2013). The nature of feelings: Evolutionary and neurobiological origins. *Nature Reviews Neuroscience*, 14(2), 143-152. <https://doi.org/10.1038/nrn3403>
- Danón, L. (2021). Conceptos en animales no humanos. In *Enciclopedia de la Sociedad Española de Filosofía Analítica*. <http://www.sefaweb.es/conceptos-en-animales-no-humanos>
- Dantzer, R., Arnone, M., & Mormede, P. (1980). Effects of frustration on behaviour and plasma corticosteroid levels in pigs. *Physiology & Behavior*, 24(1), 1-4. [https://doi.org/10.1016/0031-9384\(80\)90005-0](https://doi.org/10.1016/0031-9384(80)90005-0)
- De Waal, F. B. M. (1991). The chimpanzee's sense of social regularity and its relation to the human sense of justice. *American Behavioral Scientist*, 34(3), 335-349. <https://doi.org/10.1177/0002764291034003005>
- De Waal, F. B. M. (2000). Attitudinal reciprocity in food sharing among brown capuchin monkeys. *Animal Behaviour*, 60(2), 253-261. <https://doi.org/10.1006/anbe.2000.1471>
- Dubreuil, D., Gentile, M., & Visalberghi, E. (2006). Are Capuchin Monkeys (*Cebus Apella*) Inequity Averse? *Proceedings of the Royal Society B: Biological Sciences*, 273, 1223-1228. <https://doi.org/10.1098/rspb.2005.3433>
- Dupré, J. (2001). *Human nature and the limits of science*. Oxford: Oxford University Press. <https://doi.org/10.1093/0199248060.001.0001>
- Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223-241. <https://doi.org/10.1177/1745691612460685>

- Feldman Barrett, L. (2017). *How emotions are made: The secret life of the brain*. Boston: Houghton Mifflin Harcourt.
- Feldman Barrett, L., & Russell, J. A. (Eds.). (2015). *The psychological construction of emotion*. New York: The Guilford Press.
- Fiske, A. P., & Rai, T. S. (2015). *Virtuous Violence: Hurting and Killing to Create, Sustain, End, and Honor Social Relationships*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781316104668>
- Fletcher, G. E. (2008). Attending to the outcome of others: Disadvantageous inequity aversion in male capuchin monkeys (*Cebus apella*). *American Journal of Primatology*, 70(9), 901-905. <https://doi.org/10.1002/ajp.20576>
- Fox, P. T., & Friston, K. J. (2012). Distributed processing; distributed functions? *NeuroImage*, 61(2), 407-426. <https://doi.org/10.1016/j.neuroimage.2011.12.051>
- Freidin, E., & Mustaca, A. E. (2004). Frustration and sexual behavior in male rats. *Animal Learning & Behavior*, 32(3), 311-320. <https://doi.org/10.3758/BF03196030>
- George, N., & Sunny, M. M. (2019). Challenges to the Modularity Thesis Under the Bayesian Brain Models. *Frontiers in Human Neuroscience*, 13, 353. <https://doi.org/10.3389/fnhum.2019.00353>
- Gilby, I. C. (2006). Meat sharing among the Gombe chimpanzees: Harassment and reciprocal exchange. *Animal Behaviour*, 71(4), 953-963. <https://doi.org/10.1016/j.anbehav.2005.09.009>
- Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, 276(1657), 699-706. <https://doi.org/10.1098/rspb.2008.1324>
- Gould, S. J. (1996). *The mismeasure of man* (Rev. and expanded). New York: Norton.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814-834. <https://doi.org/10.1037//0033-295X.108.4.814>
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage Books.
- Hemelrijk, C. K. (1997). Reciprocation in apes: From complex cognition to self-structuring. En W. C. McGrew, L. F. Marchant, & T. Nishida (Eds.), *Great Ape Societies* (pp. 185-195). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511752414.016>
- Hodgkinson, G. P., Langan-Fox, J., & Sadler-Smith, E. (2008). Intuition: A fundamental bridging construct in the behavioural sciences. *British Journal of Psychology*, 99, 1-27. <https://doi.org/10.1348/000712607X216666>




- Jakovcevic, A., Elgier, A. M., Mustaca, A. E., & Bentosela, M. (2013). Frustration behaviors in domestic dogs. *Journal of Applied Animal Welfare Science*, 16(1), 19-34. <https://doi.org/10.1080/10888705.2013.740974>
- Kahneman, D. (2012). *Pensar rápido, pensar despacio*. Buenos Aires: Debate.
- Korsgaard, C. M. (2006). Morality and the Distinctiveness of Human Action. En S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 98-119). Princeton: Princeton University Press. <https://doi.org/10.1515/9781400830336-008>
- Levins, R., & Lewontin, R. C. (2009). *The dialectical biologist*. Delhi: Aaker Books.
- McPeake, K. J., Collins, L. M., Zulch, H., & Mills, D. S. (2021). Behavioural and Physiological Correlates of the Canine Frustration Questionnaire. *Animals*, 11(12), 3346. <https://doi.org/10.3390/ani11123346>
- Mercier, H., & Sperber, D. (2017). *The Enigma of Reason: A New Theory of Human Understanding*. Cambridge: Harvard University Press. <https://doi.org/10.4159/9780674977860>
- Mosterín, J. (2011). *La Naturaleza humana*. Madrid: Espasa.
- Palecek, M. (2017). Modularity of Mind: Is It Time to Abandon This Ship? *Philosophy of the Social Sciences*, 47(2), 132-144. <https://doi.org/10/gm5rzb>
- Papini, M. R. (2003). Comparative Psychology of Surprising Nonreward. *Brain, Behavior and Evolution*, 62(2), 83-95. <https://doi.org/10.1159/000072439>
- Papini, M. R., Penagos-Corzo, J. C., & Pérez-Acosta, A. M. (2019). Avian Emotions: Comparative Perspectives on Fear and Frustration. *Frontiers in Psychology*, 9, 1-14. <https://doi.org/10.3389/fpsyg.2018.02707>
- Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*. New York: Penguin.
- Prinz, J. J. (2006). Is the mind really modular? En R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 22-36). Malden: Blackwell.
- Raihani, N. J., Thornton, A., & Bshary, R. (2012). Punishment and cooperation in nature. *Trends in Ecology & Evolution*, 27(5), 288-295. <https://doi.org/10.1016/j.tree.2011.12.004>
- Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2012). No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences*, 109(37), 14824-14829. <https://doi.org/10.1073/pnas.1203179109>
- Roma, P. G., Silberberg, A., Ruggiero, A. M., & Suomi, S. J. (2006). Capuchin monkeys, inequity aversion, and the frustration effect. *Journal of Comparative Psychology*, 120(1), 67-73. <https://doi.org/10.1037/0735-7036.120.1.67>

- Sahlins, M. (2008). *The Western illusion of human nature*. Chicago: Prickly Paradigm Press.
- Schino, G., & Aureli, F. (2009). Reciprocal altruism in primates. Partner choice, cognition, and emotions. *Advances in the Study of Behavior*, 39, 45-69. [https://doi.org/10.1016/S0065-3454\(09\)39002-6](https://doi.org/10.1016/S0065-3454(09)39002-6)
- Scotto, S. C. (2022). Cognición moral y cognición psicológica: Las intuiciones vienen primero. *Revista de Humanidades de Valparaíso*, 19, 15-42. <https://doi.org/10.22370/rhv2022iss19pp15-42>
- Silberberg, A., Crescimbeni, L., Addessi, E., Anderson, J. R., & Visalberghi, E. (2009). Does inequity aversion depend on a frustration effect? A test with capuchin monkeys (*Cebus apella*). *Animal Cognition*, 12(3), 505-509. <https://doi.org/10.1007/s10071-009-0211-6>
- Solomon, R. C. (1995). Justice as vengeance, vengeance as justice. A partial defense of Polymarchus. En J. P. Sterba (Ed.), *Morality and Social Justice: Point/counterpoint*. Maryland: Rowman & Littlefield.
- Sperber, D. (1994). The modularity of thought and the epidemiology of representations. En L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind* (pp. 39-67). Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9780511752902.003>
- Sperber, D., & Baumard, N. (2012). Moral Reputation: An Evolutionary and Cognitive Perspective: Moral Reputation. *Mind & Language*, 27(5), 495-518. <https://doi.org/10.1111/mila.12000>
- Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences*, 8(2), 60-65. <https://doi.org/10.1016/j.tics.2003.12.003>
- Stout, S. C., Boughner, R. L., & Papini, M. R. (2003). Reexamining the frustration effect in rats: Aftereffects of surprising reinforcement and non-reinforcement. *Learning and Motivation*, 34(4), 437-456. [https://doi.org/10.1016/S0023-9690\(03\)00038-9](https://doi.org/10.1016/S0023-9690(03)00038-9)
- Suhler, C. L., & Churchland, P. (2011). Can Innate, modular «foundations» explain morality? Challenges for Haidt's Moral Foundations Theory. *Journal of Cognitive Neuroscience*, 23(9), 2103-2116; discussion 2117-2122. <https://doi.org/10/fsr92d>
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35-57. <https://doi.org/10.1086/406755>
- Wilson, J. Q. (1997). *The moral sense*. New York: Free Press.
- Wilton, R. N., Strongman, K. T., & Nerenberg, A. (1969). Some Effects of Frustration in a Free Responding Operant Situation. *Quarterly Journal of Experimental Psychology*, 21(4), 367-380. <https://doi.org/10.1080/14640746908400232>

eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.31172>

## FRANS DE WAAL, FILÓSOFO DO DIREITO

### *Frans de Waal, Legal Philosopher*

Fábio PORTELA LOPES DE ALMEIDA\*   
*Tribunal Superior do Trabalho, Brasília, Distrito Federal, Brasil*  
fabio.portela@gmail.com

Recibido: 25/01/2023    Revisado: 13/03/2023    Aceptado: 11/08/2023

**RESUMO:** A alegação de que apenas os humanos são capazes de raciocinar e agir normativamente é profundamente inconsistente com a teoria evolutiva, embora estejam na raiz de muitas premissas assumidas como incontestáveis na teoria jurídica. Recentemente, nos últimos trinta ou quarenta anos, a etologia e a primatologia vêm apresentando evidências de que outros animais também são agentes capazes de cooperar com base em normas sociais. Nesse âmbito, as pesquisas de Frans de Waal ofereceram um pilar fundamental à investigação do comportamento de primatas e outros mamíferos, como elefantes e golfinhos. O presente artigo pretende apresentar como a obra do etólogo pode fornecer subsídios à revisitação de temas clássicos na teoria do direito. No texto, discute-se como Frans de Waal possibilita incorporar na filosofia jurídica a adequada compreensão da cognição moral como lastro à teorização sobre a noção de sujeito de direito e os direitos dos animais, bem como

\* Doutor e Mestre em Direito (FD/UnB) e Mestre em Filosofia (FIL/UnB). Visiting Researcher em pesquisa doutoral na Harvard Law School (2013/14) e em pesquisa pós-doutoral na Europa-Universität Flensburg (2019). Pesquisador dos Grupos de Pesquisa DISCO (UnB) e GeNe (UFMG). Autor dos livros *Constitution: the Darwinian Evolution of a Societal Structure* (Nomos, 2020) e *Liberalismo Político, Constitucionalismo e Democracia: a Questão do Ensino Religioso nas Escolas Públicas* (Argumentum, 2007).

arquitetar um contra-argumento à falácia naturalista como obstáculo a uma visão naturalizada do direito.

*Palavras-chave:* teoria evolutiva, Frans de Waal, cognição moral, direitos dos animais, falácia naturalista.

**ABSTRACT:** The claim that only humans are able to reason and act normatively is deeply inconsistent with evolutionary theory, although it is at the root of many premises assumed to be undisputed in legal theory. Recently, in the last thirty or forty years, ethology and primatology have presented evidence that other animals are also agents capable of cooperating based on social norms. In this context, Frans de Waal's research offered a fundamental pillar for the investigation of the behavior of primates and other mammals, such as elephants and dolphins. This article intends to present how the work of the ethologist can provide subsidies to the revisitation of classical themes in the theory of law. The text discusses how Frans de Waal makes it possible to incorporate into legal philosophy the adequate understanding of moral cognition as a ground to theorize about the notion of legal subject and animal rights, as well as to architect a counterargument to the naturalistic fallacy as an obstacle to a naturalized conception of law.

*Keywords:* evolutionary theory, Frans de Waal, moral cognition, animal rights, naturalistic fallacy.

## 1. INTRODUÇÃO

A teoria jurídica assume implicitamente o pressuposto de que o direito é inerentemente limitado ao *Homo sapiens* a única espécie animal capaz de raciocinar normativamente. Como resultado, a moralidade e o direito só existem em nosso mundo. Esta afirmação não é tão trivial como poderia soar à primeira vista, no entanto, pois assume muitas suposições a respeito da cognição moral. Nossa capacidade de entender normas sociais, aplicá-las e nos rebelarmos contra situações injustas em nosso relacionamento diário com os outros, são assumidas como uma capacidade cognitiva humana obviamente derivada da racionalidade.

No entanto, a alegação de que apenas os humanos são capazes de raciocinar normativamente é profundamente inconsistente com a teoria evolutiva. A partir de uma abordagem darwinista, seria altamente improvável que uma única espécie na natureza exibisse traços psicológicos tão complexos não derivados da estrutura cognitiva presente em seus ancestrais.

A evolução de estruturas complexas é o produto cumulativo de processos incrementais e contínuos, não da ruptura dramática. Como resultado, o surgimento de um traço psicológico tão complexo como a cognição normativa dificilmente poderia evoluir tão abruptamente, nos 200.000 anos de existência de nossa espécie. Uma suposição mais plausível é a de que nossa psicologia social foi gradualmente moldada pela evolução ao longo de milhões de anos e, como resultado, compartilhamos uma enorme proporção de nossa cognição normativa com nossos ancestrais.

Pesquisas recentes têm confirmado essa assertiva, trazendo muitas questões sobre os fundamentos do direito. Tradicionalmente, os estudiosos jurídicos assumem como um dado certas preconcepções sobre a legitimidade do direito e seus fundamentos morais, usualmente fundadas - ainda que nem sempre explicitamente - na natureza racional da espécie humana. Nos manuais, os fundamentos do direito são usualmente apresentados como o mero resultado da história institucional ou de uma deliberação em um suposto "estado de natureza" típico da filosofia contratualista desenvolvida na modernidade. A capacidade de viver em um ambiente social estruturado a partir de normas jurídicas, éticas e morais, é sempre assumida como um pressuposto não problemático e exclusivo de nossa espécie.

Assumir a competência normativa como uma habilidade inexplicada de nossa condição humana, contudo, esconde muitas questões relevantes sobre o próprio papel do direito. Por que as comunidades humanas estão sempre organizadas em torno de certos princípios normativos? A capacidade de cumprir normas e avaliar o descumprimento de padrões normativos surgiu em outros primatas? Se não surgiu, como podemos explicar a singularidade de nossa espécie do ponto de vista evolutivo? Por outro lado, se essa capacidade também está presente em outros primatas, seríamos nós, humanos, tão diferentes deles?

A teoria jurídica tradicional não diz nada sobre esses tópicos fundamentais.

Recentemente, nos últimos trinta ou quarenta anos, desenvolveu-se um enorme conjunto de pesquisas interdisciplinares sobre esses temas. Sabemos mais sobre como a cooperação baseada em normas comportamentais evoluiu na natureza e, particularmente, como ela depende de muitas premissas psicológicas que são apenas timidamente exploradas na teoria jurídica. Ao assumir que somos capazes de ler textos legais, interpretá-los e aplicar normas normativas a situações cotidianas, a teoria jurídica levanta uma enorme questão sobre as origens do direito: como, em primeiro lugar, nós humanos nos tornamos capazes de tarefas cognitivas tão complexas?

Muitos campos interdisciplinares começaram a abordar essas questões nas últimas décadas. Antropólogos, biólogos, economistas, arqueólogos, psicólogos, sociólogos, neurologistas e primatólogos lançaram luz sobre como nossa competência moral emergiu na natureza e no curso de nossa história como espécie na Terra.

Um dos nomes proeminentes neste campo de pesquisa relativamente novo sobre as origens da moralidade e do comportamento normativo é Frans de Waal. Mais do que isso, a contribuição do primatólogo holandês é primordial para uma compreensão cientificamente informada sobre o direito. Seu trabalho com chimpanzés, gorilas e bonobos mostra que esses primatas exibem um complexo conjunto de comportamentos sociais. Capazes de viver em grupos sociais estáveis onde conflitos emergem diariamente sobre fatores como a hierarquia social, estabelecimento de coalizões, compartilhamento de alimentos e acesso sexual às fêmeas do grupo. Compor tais conflitos demanda uma sofisticada psicologia social, que inclui um senso de justiça, a habilidade de compreender relações de reciprocidade e se envolvem em processos de resolução de conflitos. Ao contrário do que nossos velhos livros de processo civil dizem, os humanos não inventaram os processos de autocomposição de controvérsias.

Em seu livro *The Age of Empathy*, por exemplo, Frans de Waal afirma que “nossos ancestrais têm sido especialmente bons em entender as emoções dos outros” (Waal, 2010). O primatólogo ressalta que nós humanos temos a “capacidade de ler expressões faciais e linguagem corporal” de uma maneira que nos permita prever o comportamento e as intenções dos outros. Essa faculdade de empatia permitiu que nossos ancestrais sobrevivessem em ambientes hostis onde a compreensão uns com os outros era essencial para a sobrevivência. Em nossos próprios dias, dependemos das mesmas habilidades empáticas para vivermos em paz uns com os outros. Mais do que isso, nossa capacidade cognitiva de entender e aplicar normas legais depende da empatia, uma capacidade fundamental que nos ajuda a nos envolver em interações recíprocas, entender a intenção dos outros e, finalmente, aplicar sanções sociais aos transgressores. É tal capacidade, como demonstra o trabalho do professor holandês, tem raízes em nossa ancestralidade primata.

Este artigo tem como objetivo apresentar parte da pesquisa de Frans de Waal e sua potencial contribuição para a teoria jurídica. Serão abordados, nesse sentido, vários aspectos do comportamento dos primatas (símios e macacos, principalmente) relevantes para a nossa compreensão dos fundamentos do direito e do comportamento social em geral. Impõe-se discutir, também, algumas das implicações desta pesquisa

para um teórico do direito: o que significa afirmar que os primatas não-humanos têm um senso de si mesmo? Como eles lidam com conflitos? Como interpretam e aplicam normas sociais? E como eles entendem o conceito de punição? Finalmente, algumas implicações para nossa compreensão do direito e da jurisprudência serão apresentadas, incluindo uma avaliação de percepções de Frans de Waal sobre temas como direitos dos animais, a relação entre o direito natural e o direito positivo, bem como a distinção ser/dever ser que, por décadas, transformou-se em um tabu que impediu os estudiosos jurídicos de explorar esses temas com maior seriedade.

## 2. HUMANOS, PRIMATAS: CONTEXTUALIZANDO A PESQUISA DESENVOLVIDA POR FRANS DE WAAL

Frans de Waal dedicou a maior parte de sua carreira acadêmica ao estudo do comportamento social de primatas. Já em sua tese de doutoramento, orientada pelo etólogo Jan van Hoof, demonstrou competência no estudo do comportamento de primatas, sua agressividade e capacidade de formação de alianças, a partir da investigação de macacos de Java (*Macaca fascicularis*) (Waal, Hoof & Netto, 1976).

Posteriormente, o pesquisador holandês desenvolveu um projeto de pesquisa na colônia de chimpanzés do Zoológico de Arnhem. Contrapondo-se à tese de que outros animais são irracionais, de Waal descreveu o comportamento dos primatas em termos de estratégias sociais cuidadosamente planejadas. A seu ver, somente podemos compreender adequadamente o comportamento social de outros primatas, especialmente no tocante à formação de coalizões, resolução de conflitos ou o compartilhamento de alimentos a partir da suposição de que são dotados de inteligência suficiente para compreender as diferentes estratégias. Esse projeto culminou na publicação de sua primeira grande obra, *Chimpanzee Politics* (Waal, 1989) e, mais tarde, a suposição teórica de Frans de Waal ganhou tamanha repercussão que passou a ser rotineiramente conhecida na etologia como “inteligência maquiavélica” (Byrne & Whiten, 1988).

Nas décadas seguintes, a pesquisa de Frans de Waal se dedicou ao estudo das emoções e intenções de primatas, o que muitas vezes levou a acusações de antropomorfismo do comportamento animal. De acordo com os críticos, de Waal estaria atribuindo a animais características tipicamente humanas, o que era problemático do ponto de vista metodológico. Afinal, essa abordagem poderia assumir como uma premissa teórica exatamente o que deveria ser demonstrado empiricamente - que

o comportamento animal seria explicável a partir das categorias típicas que aplicamos aos humanos (Wright, 2006). Em resposta, de Waal assumiu a própria validade epistemológica do antropomorfismo como estratégia de investigação, salientando que, de outro modo, talvez jamais fosse possível compreender adequadamente o comportamento dos primatas estudados.

Então, a resposta para a pergunta “O antropomorfismo não é perigoso?” é que, sim, é perigoso para aqueles que desejam manter um muro entre humanos e outros animais. Coloca todos os animais, incluindo os humanos, no mesmo plano explicativo. Não é perigoso, porém, para aqueles que trabalham a partir de uma perspectiva evolucionária, desde que tratem as explicações antropomórficas como hipóteses (Burghardt, 1985). O antropomorfismo é uma possibilidade entre muitas, mas deve ser levada a sério, já que aplica intuições sobre nós mesmos a criaturas muito parecidas conosco. É a aplicação do autoconhecimento humano ao comportamento animal. O que poderia estar errado com isso? Aplicamos a intuição humana em matemática e química, então por que deveríamos suprimi-la no estudo do comportamento animal? Mais forte ainda: alguém realmente acredita que o antropomorfismo é evitável (Spada, 1997)? (Waal *et al.*, 2006).

Como resultado de suas pesquisas, o primatólogo passou cada vez mais a denotar como as capacidades cognitivas humanas eram tipicamente encontradas também em outros animais de nossa linhagem primata ou mesmo em outros mamíferos. Não se trata, contudo, do mero compartilhamento pontual de determinadas habilidades. Frans de Waal sustenta que os mamíferos apresentam, em maior ou menor grau, um conjunto de traços que evoluíram progressivamente e sustentam a interdependência emocional entre indivíduos.

Entre essas habilidades, destacam-se o reconhecimento da própria identidade, o contágio emocional, a empatia cognitiva, a identificação da perspectiva do outro e a teoria da mente (Preston & Wall, 2001). Embora nem todos os mamíferos apresentem tais capacidades, habilidades socialmente complexas já foram identificadas em animais como golfinhos e os grandes primatas (orangotangos, chimpanzés, gorilas e bonobos). Tais espécies têm demonstrado habilidades como a formação de coalizões; o altruísmo, inclusive dirigido a indivíduos de outras espécies (Waal, 2010); a aversão a injustiças (Brosnan & Waal, 2003); e a resolução de conflitos com a intervenção de um terceiro (Das, 2000; Waal, 2010). Matriarcas gorilas, por exemplo, intervêm para arbitrar conflitos de terceiros com vistas a pacificar o relacionamento, reconciliando os agressores após um confronto. Mesmo baleias são capazes de demonstrar



altruísmo extremo, se interpondo entre um navio arpoador e um membro de seu grupo para protegê-lo (Waal, 2010).

Entre as evidências que sustentam a tese, Frans de Waal menciona como os bonobos (e golfinhos) utilizam as relações sexuais para compor conflitos e elefantes têm autoconsciência de si mesmos, inclusive apresentando comportamento de luto após a morte de entes queridos (Plotnik, Waal & Reiss, 2006). Além disso, como ilustra experimento conduzido por Joshua Plotnik e Frans de Waal, juntamente com outros etólogos, os paquidermes são capazes de cooperar ativamente para produzir um resultado desejado (Plotnik *et al.*, 2011).

A pesquisa de Frans de Waal também apresenta contribuições importantes para a compreensão de como opera a cognição normativa entre animais não-humanos. Experimentos elaborados e conduzidos por ele e Sarah Brosnan, por exemplo, identificaram que primatas apresentam comportamentos que seriam ainda mais próximos daquilo que costumamos chamar de justiça. Alguns experimentos projetados pelos etólogos Sarah Brosnan e Frans de Waal demonstram, por exemplo, que chimpanzés têm aversão à desigualdade (Brosnan & Waal, 2003).

Um primeiro experimento foi projetado a fim de verificar se macacos capuchinhos (*Cebus apella*) são capazes de compartilhar alimentos mesmo sem motivo algum para fazê-lo. Para testar essa hipótese, os pesquisadores posicionaram algumas nozes em um pote, isolado de dois macacos por meio de uma barreira de vidro que contém uma pequena abertura, suficiente apenas para passar um dos braços do macaco. Para abrir o pote, além disso, é preciso utilizar uma pedra, colocada do lado onde estão os macacos. É necessário ter os dois para ter acesso às nozes. A solução, claro, é passar a pedra para o outro lado para que o macaco que está com as nozes possa abrir o pote e pegá-las. Mas isso traz um dilema para o macaco que tem acesso à pedra: por que ele ajudaria o outro macaco? Que garantia ele tem que o macaco não pegará todas as nozes para si, ao invés de compartilhá-las? Apesar disso, ele confia no outro macaco e lhe entrega a pedra; e, em contrapartida, o outro pega as nozes e as divide igualmente com o primeiro. A ênfase, aqui, está no igualmente: ele poderia ter entregado apenas uma noz, ou até nenhuma, mas preferiu dividir igualmente o produto de seus esforços. A conclusão, deste e de outros estudos, é clara: os macacos têm uma mente que valoriza de algum modo a justiça e uma certa divisão igualitária de recursos.

Em sua primeira grande obra, Frans de Waal já sugeria a presença de um senso de justiça entre os chimpanzés em suas interações individuais (Waal, 1989), mas passou a abordar o tema mais consistentemente pouco depois. Em artigo de 1991, de Waal já salientava que chimpanzés

possuíam sentido de regularidade em suas relações interpessoais, que o autor reputa como fundante do próprio senso de justiça humano. Tal senso de regularidade, tal como observado nos símios, se revela na observância de normas reforçadas socialmente pela punição de transgressores (Waal, 1991, 343). Mais recentemente, de Waal sustentou que há evidências concretas de que animais se comportam normativamente – definida a normatividade como “aderência a um ideal ou padrão, é ampla a evidência de que animais tratam suas relações sociais desse modo. Em outras palavras, eles perseguem valores sociais” (Waal, 2014, 185).

Nesses termos, a obra de Frans de Waal se afirma como potencial inspiração para a discussão de inúmeros temas da teoria do direito. Na próxima seção, abordarei alguns dos grandes temas da teoria jurídica que podem ser rediscutidos e reformulados a partir da obra do primatólogo holandês, incluindo a teorização sobre as origens do direito, a distinção entre ser/dever ser, os fundamentos da normatividade na psicologia moral e social humana, bem como questões jurídicas mais tradicionais, como a possibilidade de animais serem portadores de direitos e a relação entre o direito natural e o direito positivo.

### **3. ENTRE ROUSSEAU E KELSEN: CONTRIBUIÇÕES DE FRANS DE WAAL À TEORIA JURÍDICA**

Conquanto a teorização sobre as origens do direito já tenha sido influenciada por teorias evolucionárias no final do século XIX e no início do século XX, com os inestimáveis contributos de Henry Maine (1906), John Henry Wigmore e Oliver Wendell Homes Jr (1900), aos poucos tais referências desapareceram das discussões a esse respeito. Com as notáveis exceções de F.A. Hayek (1998) e, mais recentemente, Richard Posner (1980), os livros sobre teoria e filosofia do direito sequer mencionam a existência de uma tradição no pensamento jurídico influenciada por uma abordagem naturalística a respeito das origens do direito.

Hodiernamente, tal discussão é tratada a partir de premissas bastante questionáveis do ponto de vista da antropologia, da etologia, da psicologia, da economia e das demais ciências que investigam as origens da sociabilidade humana. Influenciadas, ainda que implicitamente, por teorias contratualistas iluministas, muitas das obras mais populares afirmam o direito como o produto da racionalidade humana. Segundo essa concepção, o direito teria derivado de convenções ancestrais que, superando o passado animalesco no qual as controvérsias eram solucionadas a partir do uso da força, passou a instituir normas impositivas

de resolução dos conflitos, bem como normas de regência do comportamento humano.

Outra discussão clássica, mas que ainda é muito ensinada nas faculdades de direito, se apoia na distinção entre direito natural e direito positivo. A teoria do direito natural decorre de uma tradição antiga na teoria jurídica, que busca justificar o direito a partir de proposições morais universalmente aceitáveis, baseadas na unidade fundamental da racionalidade humana. O direito positivo – direito legislado ou imposto pelas instituições jurídico-políticas - varia em diferentes contextos político-geográficos, mas o direito natural, como expressão da racionalidade, é universal e imutável. Como afirma Daniel Chernilo, a teoria do direito natural baseia-se na crença em “um conjunto universal de leis que são válidas independentemente do lugar, do tempo ou da cultura” (Chernilo, 2013).

A ideia de que a validade do direito é apoiada em normas pré-sociais acompanhou a filosofia jurídica desde a sua criação. Dos filósofos antigos e medievais Aristóteles, Cícero e Aquino aos modernos Hobbes, Pufendorf, Locke, Kant, Rousseau e Vitória, e aos teóricos contemporâneos como Hans Kelsen, Herbert Hart, Alasdair MacIntyre e John Finnis, a maioria dos filósofos do direito escreveu sobre o assunto, seja para abraçar o direito natural ou para argumentar contra sua existência como pressuposto de validade do direito positivo (Contreras, 2013).

A contribuição de Frans de Waal a tais reflexões parece evidente. Sua pesquisa etológica apresenta evidências concretas de que o comportamento social de animais próximos, na ancestralidade evolutiva, ao *Homo sapiens*, sinaliza a pré-existência de uma psicologia inata repleta de elementos necessários à cognição moral. Tal percepção oferece um caminho único de compreensão de como determinados pressupostos cognitivos foram necessários à evolução de uma espécie como a nossa, capaz de raciocinar a partir de normas morais (e eventualmente jurídicas) e conduzir seu comportamento de acordo com elas.

Esse caminho conduz naturalmente a um importante fundamento para os direitos dos animais, na medida em que demonstra a riqueza da psicologia moral/social de outros animais – alguns nem tão próximos de nós, humanos, na escala evolutiva, como golfinhos e mesmo baleias. Esses animais, além da evidente capacidade de sentir o sofrimento.

Contudo, um obstáculo importante se apresenta a uma abordagem da distinção entre ser/dever ser, tão basilar ao raciocínio jurídico. Se a validade do direito repousa em uma análise apenas normativas sobre os direitos e deveres, como sustentado pelas teorias jusfilosóficas mais aceitas em nosso tempo, o papel da psicologia parece sepultado,

uma vez que qualquer derivação de normas a partir de características da psiquê humana aparentemente incorrer na chamada falácia naturalista. Mas Frans de Waal, novamente, oferece contribuição significativa para esse problema, ofertando um caminho lógico de solução que apresenta a conexão entre a ontologia psicológica e a deontologia das normas, sem afrontar o postulado básico da distinção jusfilosófica clássica.

### **3.1. *Da psicologia animal ao direito: descobrindo as raízes cognitivas da normatividade***

A teoria jurídica tradicional afirma que o direito é fruto da racionalidade humana (Ferraz Júnior, 2013, 157). De acordo com a leitura proposta, primitivamente os conflitos eram solucionados com base exclusiva na violência e por meio do uso da força. No mundo natural, os animais são incapazes de agir moralmente e resolverem pacificamente conflitos por não serem racionais e não viverem em um mundo cultural constituído por normas morais e jurídicas. Em razão da vida em sociedade e a cultura - continua o argumento -, nos tornamos agentes racionais e capazes de agir conforme normas que regulam nosso comportamento e de respeitar as instituições sociais.

Muitos livros introdutórios ao estudo do direito, por exemplo, afirmam que, antes de as sociedades humanas alcançarem um estágio cultural em que uma instituição como o Estado surgiu, os conflitos sociais eram resolvidos com base na luta entre indivíduos a fim de obter o objeto de seu desejo pela força. Como não havia normas para solucionar conflitos, estes eram decididos com base na violência. Em um manual popular de introdução ao direito processual civil, por exemplo, consta a seguinte passagem: “Nos primórdios da humanidade (...) imperava a lei do mais forte, em que o conflito era resolvido pelos próprios indivíduos” (Alvim, 2001, 10-11). E em outra obra, estudada como clássico, lê-se que “a fim de desfazer o conflito surge uma primeira solução – a violência. É a forma primitiva, e ainda não totalmente extinta, de solução dos conflitos de interesses individuais ou coletivos. É o predomínio da força” (Santos, 2011, 11). Afirmações como essas podem ser lidas e repetidas como verdade absoluta em alguns principais manuais de direito, que ensinam a gerações inteiras de estudantes como, graças à racionalidade, o direito e o Estado trouxeram à humanidade a capacidade de resolver pacificamente seus conflitos.

No entanto, mesmo que grande parte da população jamais tenha lido um único artigo de lei, o direito é cumprido cotidianamente sem que as pessoas recorram ao Poder Judiciário a todo instante para solucionar suas controvérsias. Mesmo sem ter lido uma única linha do Código Penal, a maioria das pessoas conduz seus atos de forma a jamais praticar os crimes moralmente mais graves. O raciocínio moral é intuitivo – e o direito, a religião e a moral efetivamente regulam a conduta humana porque somos, antes mesmo de sermos influenciados por essas instituições – capazes de agir de acordo com normas.

Mas por que somos assim? Uma das contribuições mais fundamentais e evidentes de Frans de Waal à teoria do direito diz respeito justamente à origem da capacidade cognitiva de raciocinar normativamente, tão típica do *Homo sapiens*. E uma grande lição do primatologista se traduz no reconhecimento de que outros animais também dispõem de uma certa capacidade cognitiva de responder a situações que poderiam ser descritas moralmente. Suas pesquisas sugerem que os chimpanzés adotam comportamentos especificamente voltados para a resolução de conflitos. Depois de um confronto entre dois chimpanzés, por exemplo, não é incomum que o vencedor da disputa se reconcilie com o perdedor, passando algum tempo catando seus parasitas, num processo conhecido como catação (*grooming*). Trata-se de comportamento bastante custoso em termos biológicos, na medida em que o estado de relaxamento exigido pela catação expõe os chimpanzés a predadores, além de demandar tempo considerável que poderia ser dedicado a outras atividades. De Waal também registra, entre outras condutas, a intervenção de chimpanzés de status hierárquico superior em conflitos do bando, evitando que as disputas internas sejam danosas a longo prazo para o grupo (Waal, 2000, 586-590).

Uma das características mais marcantes identificadas por Frans de Waal diz respeito à capacidade de compreender relações sociais recíprocas, uma premissa na construção cognitiva de um certo senso normativo (de justiça) entre primatas, orientado pela aversão à desigualdade. Trabalhando com macacos capuchinhos (em experimento que mais tarde foi repetido com chimpanzés, com resultados similares), de Waal e Sarah Brosnan elaboraram a seguinte experiência: testando dois macacos por vez, Sarah ofereceu uma ficha a cada um deles, que poderia ser trocada por comida. Nas primeiras experiências, Sarah trocava a pedra por pepinos, e os macacos efetuavam a troca sem qualquer conflito. Mas, quando Sarah tornou a relação de troca sutilmente mais desigual, passando a oferecer uvas para um dos macacos e continuando a oferecer pepinos para o outro – ciente de que a espécie prefere uvas a pepinos - o comportamento

do macaco que continuou a receber o pepino mudou drasticamente. Ao invés de permanecer satisfeito com o pepino, o macaco prejudicado sinalizou a injustiça da situação, tornando-se agitado e agressivo e chegando a jogar as pedras, ou mesmo os pepinos, para fora da jaula.

De acordo com a interpretação de Frans de Waal, isso ocorreu como reação à situação de injustiça (Waal, 2010, 187). O fato de o macaco capuchinho reconhecer seu prejuízo comparativo indicaria a capacidade de identificar e reagir à desigualdade, consistindo em um indício das origens evolutivas do senso de justiça (Waal, 2009, 187-189). Pesquisas mais recentes, realizadas com chimpanzés, indicam que esses símios são capazes de reagir à desigualdade de tratamento também em relação a outros indivíduos. Modificando o experimento, Brosnan *et al.* (2010a) treinaram chimpanzés para trocar pequenos objetos por comida, e testaram como pares de chimpanzés reagiam a vários níveis de recompensa. Como nos experimentos originais, os chimpanzés rejeitavam cenoura quando percebiam que o outro símio era recompensado com uvas. Contudo, distintamente, também era elevada a taxa de rejeição da recompensa quando o próprio indivíduo ganhava uva e o outro chimpanzé, cenouras. Na percepção dos autores do experimento, esse comportamento revela o monitoramento ativo e a reação a diferenças em desfavor de outros (Brosnan *et al.*, 2010b).

A lógica evolutiva da reação à desigualdade nos níveis individual e social é explicada como uma adaptação esperada em animais sociais complexos como os grandes símios e, além disso, é uma das fundações psicológicas de sistemas mais complexos de cooperação social. Nas palavras do cientista:

De onde essas reações vêm? Elas provavelmente evoluíram a serviço da cooperação. A preocupação com o que os outros ganham pode parecer de pouca importância e irracional, mas no longo prazo impede alguém de ser passado para trás. É do interesse de todos desencorajar a exploração e os oportunistas (*free-riding*), e ter certeza de que os interesses de cada um são levados a sério (Waal, 2010, 187-188).

Pesquisas como essas indicam que, contrariamente ao senso comum teórico repetido em manuais de estudos jurídicos, formas de resolução de conflitos e de identificação e reação a situações sociais injustas a partir de padrões normativos prévios não são uma exclusividade humana. Muito antes de formas sociais de organização centralizadas no poder coercitivo do Estado emergirem, a seleção natural moldou a psicologia de animais sociais de modo a que se tornassem cognitivamente capazes de solucionar suas controvérsias ativamente. Se chimpanzés e outros

grandes primatas são próximos do *Homo sapiens* na escala evolutiva, comportamentos como a punição social de transgressores e a reconciliação entre agressor e ofendido, entre outros investimentos cooperativos, já foram identificados em outras espécies animais socialmente complexas, como golfinhos, morcegos vampiros e elefantes, denotando a lógica evolutiva de estratégias adaptativas convergentes (Carter, 2014).

De acordo com Giuseppe Lorini (2018), tradicionalmente se considera a linguagem como pré-condição da ação nômica e da normatividade; contudo, esta é uma premissa a ser investigada empiricamente. De acordo com ele, as pesquisas mais recentes em etologia vêm mostrando que a resposta é negativa, na medida em que diversos animais apresentam comportamentos que apenas podem ser explicados mediante a suposição de que compreendem e aplicam regras. Assim, em sua perspectiva, a linguagem não é um requisito à normatividade.

Por outro lado, pesquisadores como o jurista italiano Rodolfo Sacco (1995) têm sustentado a existência de normas em outras espécies animais. De acordo com ele, há um certo viés em não reconhecer a normatividade em outras espécies por viverem em sociedades não linguísticas. Mas, quando concebemos o direito na pré-história, em sociedades ainda pré- ou proto-linguísticas, não é difícil imaginar a presença de regras, inobstante inexista a correspondente descrição linguística. O registro arqueológico, afinal, consigna a presença de ferramentas elaboradas a partir de ossadas ou pedras lascadas, em número suficiente para sugerir que tais instrumentos tinham relevância social – tanto para construí-las quanto para coordenar sua boa utilização em comunidade. Sacco opina, ainda, que muitos atos de natureza normativa eram praticados na ausência de linguagem, a exemplo da marcação de território, de regras de acesso a relacionamentos sexuais, deveres relativos a filhos e parentes, bem como inúmeros outros exemplos relativos ao cumprimento de obrigações recíprocas e observância de tabus. Ainda que não haja normas escritas, e ainda que espontaneamente, os indivíduos relevantes aderem a elas, revelando não apenas a sua existência social, mas também sua própria validade. É inequívoca a existência de normas subjacentes ao comportamento, constituindo uma normatividade pressuposta na psicologia social que se revela em um direito ‘mudo’. Por outro lado, pesquisadores como o jurista italiano Sacco (1995) têm sustentado a existência de normas em outras espécies animais. De acordo com ele, há um certo viés em não reconhecer a normatividade em outras espécies por viverem em sociedades não linguísticas. Mas, quando concebemos o direito na pré-história, em sociedades ainda pré- ou proto-linguísticas, não é difícil imaginar a presença de regras, inobstante inexista a correspondente

descrição linguística. O registro arqueológico, afinal, consigna a presença de ferramentas elaboradas a partir de ossadas ou pedras lascadas, em número suficiente para sugerir que tais instrumentos tinham relevância social – tanto para construí-las quanto para coordenar sua boa utilização em comunidade. Sacco opina, ainda, que muitos atos de natureza normativa eram praticados na ausência de linguagem, a exemplo da marcação de território, de regras de acesso a relacionamentos sexuais, deveres relativos a filhos e parentes, bem como inúmeros outros exemplos relativos ao cumprimento de obrigações recíprocas e observância de tabus. Ainda que não haja normas escritas, e ainda que espontaneamente, os indivíduos relevantes aderem a elas, revelando não apenas a sua existência social, mas também sua própria validade. É inequívoca a existência de normas subjacentes ao comportamento, constituindo uma normatividade pressuposta na psicologia social que se revela em um direito ‘mudo’.

A própria posição de Frans de Waal é clara a esse respeito. É certo que, por décadas, o primatólogo holandês estudou profundamente o comportamento social de chimpanzés, tendo escrito a respeito da existência de normas sociais em sociedades de primatas não-humanos. De acordo com ele, há evidências consistentes de que os animais aderem a normas sociais, observando-as como padrões a serem cumpridos de forma a perseguir concretamente determinados valores sociais. Contudo, a natureza dessas normas ainda é pré-linguística; não derivam de racionalidade cognitiva, mas do algoritmo darwinista. Mediante a seleção natural, procedendo ao longo dos milênios, normas sociais como as mencionadas impregnaram a cognição social de diversas espécies animais, determinando, geneticamente, o escopo dos comportamentos sociais esperados. Ainda que os animais possam observar certos padrões de comportamento, isso não significa que eles sejam capazes de racionalizá-los e compreendê-los de forma clara. Têm competências sociais, mas não a capacidade de compreendê-las linguisticamente.

É nesse sentido que venho utilizando o termo ‘mente normativa’ para caracterizar a capacidade cognitiva de compreender conteúdos normativos e se comportar de acordo com expectativas deonticas (Almeida, 2011; 2013; 2020). Tal como Chomsky argumenta em relação à linguagem, processos seletivos insculpiram uma gramática moral universal a partir de critérios evolutivos próprios, que respondem a problemas biológicos que emergem de contextos cooperativos, capacitando indivíduos a operarem de acordo com os postulados da seleção por parentesco e do altruísmo recíproco, bem como, no caso humano, dos instintos sociais tribais (Jordan *et al.*, 2013; Nowak, 2006; Richerson & Boyd, 2008; Silk, 2009; Soltis, Boyd & Richerson, 1995).



A tese de Frans de Waal fornece subsídios concretos a essa tese, revelando não apenas que outros animais têm capacidades como as mencionadas – exceto no que diz respeito aos instintos sociais tribais, típicos do *Homo sapiens* –, mas também que observam normas sociais específicas que estão entranhadas em sua psicologia. A essas normas venho denominando de ‘direito natural’, em referência à menção, na tradição filosófica clássica, de normas universais e imutáveis que derivariam da própria natureza (o cosmos grego) ou, no contexto moderno, da natureza racional humana. Na presente abordagem, contudo, trata-se de um direito natural naturalizado, cujo conteúdo está impregnado em nossa própria racionalidade normativa. Ao analisarmos situações que evocam a racionalidade moral, nossa cognição já as avalia conforme modelos heurísticos com conteúdo normativo previamente incorporado à nossa psicologia ao longo da evolução de nossa ancestralidade humana.

### 3.2. *Frans de Waal e a questão dos direitos dos animais*

A defesa dos direitos dos animais é um ponto bastante polêmico na obra de Frans de Waal. O autor tem sido de extrema relevância para a causa dos direitos dos animais, na medida em que tem contribuído para a evolução do pensamento neste campo para além das considerações usuais na área do direito ambiental. Tradicionalmente, argumentos jurídicos de proteção ao bem-estar animal têm sido fundamentados na função mediata dos animais frente ao próprio bem-estar humano. Contudo, de Waal inovou no argumento, por demonstrar que os animais possuem uma natureza cognitiva complexa e, como tal, seu bem-estar é repleto de nuances que precisam ser compreendidas para prover sua adequada preservação.

No entanto, na obra *Primates and Philosophers* (Waal *et al.*, 2006), o primatólogo argumenta que as evidências demonstram que os animais possuem consciência e sentem emoções como gratidão e indignação, bem como a capacidade de sentir dor e sofrimento, é desnecessário atribuir a eles direitos. De Waal menciona inclusive o comportamento político dos chimpanzés, que negociam em suas relações sociais a formação de coalizões para elevar sua posição na hierarquia do grupo. Inobstante tais evidências fáticas devam ser consideradas ao se discutir o tratamento destinado a outras espécies – argumenta de Waal –, o *Homo sapiens* ostenta posição singular que torna criticável a atribuição de direitos a outras espécies.

No livro, de Waal critica ostensivamente a posição do professor Steven M. Wise, de Harvard, para quem os animais não são apenas objeto da propriedade humana, mas merecem os mesmos direitos incontestáveis que os seres humanos. Mencionando como absurda a posição de defensores de direitos dos animais que postulam a concessão de direitos absolutos à integridade corporal e à liberdade, de Waal sustenta que a capacidade de ter direitos exige a assunção de responsabilidades por seu titular. Ser sujeito de direitos, em sua visão, implica a capacidade de compreender seu significado e reivindicar sua proteção. Por esse motivo, inclusive, critica os que comparam o movimento pelos direitos dos animais com a abolição da escravatura como moralmente falha, na medida em que os negros escravizados foram capazes de se tornarem membros plenos da sociedade, enquanto os animais jamais o serão.

Com base nesses fundamentos, e aludindo à circunstância de que a concessão de direitos aos animais depende inteiramente da nossa boa vontade – e não da reivindicação por parte de quem os exerceria –, Frans de Waal propõe que, ao invés de nos concentrarmos nos direitos, deveríamos defender uma ética do cuidado que leve em conta a nova compreensão da vida mental e emocional dos animais. Devemos ensinar as pessoas a respeitar os animais e a estar cientes de seus interesses, com vistas a garantir as condições necessárias para assegurar seu bem-estar.

Em seção destinada a comentários por outros filósofos, Peter Singer apresenta importante contraponto à posição do primatólogo. O professor australiano é conhecido entre os ativistas dos direitos dos animais por sua obra *Animal Liberation* (2009[1975]) e pela criação, em conjunto com Paola Cavalieri, do *The Great Ape Project*, iniciativa voltada à preservação dos grandes primatas. De início, Singer reconhece que ambos adotam premissas em comum, porquanto atribuem aos animais a capacidade de sentir dor e emoções bastante significativas, rejeitando eventual interpretação antropomórfica de seu comportamento. Além disso, também perfilham o comprometimento com a prevenção do abuso de animais – afirmando, inclusive, que, “se todos compartilhassem os pontos de vista de Waal, o movimento animal alcançaria rapidamente seu objetivo mais importante” (Waal *et al.*, 2006, 153, tradução livre).

No entanto, Singer entende que a atribuição de direitos aos animais constitui importante pauta de natureza política. Reconhecer direitos básicos a todos os animais pode impulsionar a propositura de projetos de lei visando resguardar seu bem-estar, elevando os padrões de tratamento contra abuso e assegurando adequados padrões de alimentação. Menciona, como exemplo, como a legislação britânica proibitiva da

manutenção de porcos em caixas apertadas, elevou a qualidade de vida usufruída por centenas de milhares dos suínos (Waal *et al.*, 2006).

O professor australiano também enfrenta o argumento de que não devemos atribuir direitos aos animais em razão de eles serem incapazes de invocá-los, exercê-los e reivindicá-los. Contra o argumento de que é indevida a analogia entre as penúrias vivenciadas pelos animais e a escravidão porque os escravos tinham a capacidade de se tornar cidadãos plenos e aptos a participar da vida democrática em igualdade de condições, Singer aponta a existência de muitos agentes incapazes de exercer plenamente seus direitos, mas os têm reconhecidos a si. Como exemplo, cita a situação jurídica de bebês e pessoas com deficiência intelectual severa (Waal *et al.*, 2006, 154) Embora incapazes de reivindicar direitos, tais pessoas têm *status* jurídico de igualdade em relação aos demais cidadãos e, embora não possam votar e exercer determinados direitos, nem por isso são menos protegidas como portadores dos direitos de que são detentores por sua condição humana. Do mesmo modo, determinados direitos poderiam ser assegurados aos animais.

No fim das contas, ambos concordam quanto à maioria das premissas da discussão. A posição do primatólogo quanto aos direitos dos animais não diminui em nada seu trabalho como ativista pela preservação das espécies e pela melhoria significativa do tratamento dirigido aos animais pela indústria farmacêutica e cosmética, fazendas e zoológicos. De acordo com sua perspectiva, contudo, essa é uma obrigação ética que impõe responsabilidade no trato com animais equipados com uma psicologia complexa, que sofrem com os maus tratos e sentem emoções multifacetadas. Não se trata, nessa concepção, de um direito de qualquer animal que possa ser reivindicado juridicamente, mas de uma responsabilidade moral de nós, que somos capazes de compreendê-los como objeto de nossa consideração moral.

### 3.3. *Frans de Waal e a falácia naturalista*

A falácia naturalista é uma falácia lógica que decorre da derivação de juízos morais a partir de uma situação de fato que, por si só, não ostenta qualquer *status* moral. A ideia básica é que, embora seja possível observar fatos e elaborar generalizações a partir deles, não se pode logicamente derivar uma obrigação moral dessas observações. Por exemplo, não se pode concluir logicamente que se deve fazer algo simplesmente porque é o caso de que algo é o caso. Ao incorrer na falácia naturalista, padrões morais são justificados e construídos a partir de premissas factuais. A

falácia consistiria em proceder a um salto argumentativo, confundindo os domínios lógicos do descritivo com o do prescritivo. Juízos morais e normativos devem ser justificados com base em outras premissas normativas, não com fundamento em descrições ontológicas. Exemplos comuns da falácia incluem argumentar que algo está errado por não ser natural, ou justificar uma ação por ela estar de acordo com as leis da natureza – como ocorre quando alguém critica o casamento homoafetivo por ser contrário às leis da natureza.

David Hume é considerado como um dos principais precursores da falácia naturalista. Os próprios escritos de Hume sobre o assunto são limitados, mas são altamente esclarecedores em termos de compreensão de seus pensamentos sobre o assunto. Em sua obra seminal, *A Treatise of Human Nature*, Hume argumenta que os valores morais não são derivados do mundo natural ou da razão humana, mas sim originários de nossos próprios “sentimentos internos de aprovação e desaprovação”. Para Hume, o que é moralmente bom ou ruim é determinado por uma avaliação subjetiva de certos sentimentos ou emoções. Ele argumenta que os princípios morais não surgem de uma racionalidade pura, mas sim das nossas experiências sociais e da nossa natureza humana. Para Hume, os sentimentos morais são aqueles que nos levam a aprovar ou desaprovar determinadas ações. Assim, se a ação é moralmente boa, sentimos aprovação, e se é moralmente má, desaprovação. Hume acredita que os sentimentos morais são inatos e não podem ser explicados pela razão, e que eles são o único fundamento para a moralidade.

Outra contribuição significativa de Hume consiste na distinção lógica entre ser e dever ser (*is-ought problem*), também conhecida como Lei de Hume ou Guilhotina de Hume. Hume concluiu a famosa seção do tratado que se opunha ao racionalismo moral apontando que outras filosofias morais derivam conclusões normativas, ligadas ao domínio do dever ser (como o direito e a moral), a partir de premissas ontológicas (o domínio do ser) sem observar o salto lógico. Hume argumentou que essa transição é logicamente insustentável. De acordo com a distinção proposta, não podemos confundir os planos do “ser” e do “dever ser”, já que ambos os juízos pertencem a domínios lógicos distintos. Como consequência, é indevido derivar uma prescrição ou uma obrigação moral de uma declaração de fato. De acordo com o próprio Hume:

Em todo sistema de moral que até hoje encontrei, sempre notei que o autor segue durante algum tempo o modo comum de raciocinar, estabelecendo a existência de Deus, ou fazendo observações a respeito dos assuntos humanos, quando, de repente, surpreendo-me ao ver que, em vez das cópulas preposicionais usuais, como é e não é, não encontro uma só

proposição que não esteja conectada a outra por um deve ou não deve. Essa mudança é imperceptível, porém da maior importância. Pois, como esse deve ou não deve expressa uma nova relação ou afirmação, esta precisaria ser notada e explicada; ao mesmo tempo, seria preciso que se desse uma razão para algo que parece inteiramente inconcebível, ou seja, como essa nova relação pode ser deduzida de outras inteiramente diferentes. (Hume, 2009, 509 [1739-1740])

O termo 'falácia naturalista', no entanto, foi proposto pela primeira vez pelo filósofo britânico G.E. Moore em sua obra de 1903 *Principia Ethica* (Nuccetelli, 2021, 116). Na obra, Moore apresentou três princípios sobre o tema: a) o "bem" expressa o único termo ético que não admite uma análise reducionista; b) nenhum termo ético pode ser reduzido a termos puramente naturalísticos ou metafísicos; e c) nenhuma proposição moral decorre apenas de proposições não-morais (Moore, 1993[1903]). Em síntese, julgamentos morais não podem ser derivados de declarações de fato; *explicações não justificam ações*.

O projeto de uma ética evolucionista ou naturalista encontra nessas apresentações da falácia um obstáculo importante a suas pretensões. Como discutido na seção 2.1., a obra de Frans de Waal apresenta uma estreita relação entre a moral e a compreensão da evolução humana dentro de um universo de descendência de outras espécies que também apresentavam traços cognitivos necessários à razão prática. A racionalidade normativa humana, nessa visão, depende de 'blocos de construção' selecionados progressivamente no curso do processo evolutivo e, em particular, dos últimos 5 milhões de anos, quando a ancestralidade do *Homo sapiens* lentamente divergiu de outras espécies de primatas.

De acordo com os defensores da ética evolucionista, a teoria da evolução pode esclarecer aspectos sobre a natureza da moralidade, sendo capaz, inclusive, de fornecer justificção para algumas de nossas normas morais. Como sustenta Karla Chediak, contudo, essa concepção ética tem recebido críticas profundas, em parte porque a moral tem sido considerada, por muito tempo, como "a expressão máxima da indeterminação e da independência do homem em relação ao resto da natureza" (Chediak, 2006, 148). Além disso, a filósofa da biologia também denota a própria falácia naturalista como empecilho à ética evolucionista, porquanto uma visão evolutiva da moral poderia *justificar* comportamentos considerados adaptativos na biologia. Nas palavras da autora:

Se realmente é possível sustentar-se que algumas capacidades motivacionais, comportamentais e cognitivas relacionadas com a moral foram formadas ao longo da evolução da espécie humana, então, pode-se afirmar que elas sustentam, ao menos em parte, os valores humanos e

pode-se também apelar a eles para justificar as ações. Transforma-se, assim, o valor adaptativo em termos de sobrevivência e reprodução em valor moral básico (Chediak, 2006, 152).

Os estudos de Frans de Waal apresentam importantes contribuições à ética evolucionista, diante da apresentação de evidências empíricas sobre o comportamento moral de outros animais e, particularmente, dos grandes primatas. Suas observações etológicas expõem como, além de nós, outras espécies são capazes de sentir empatia, se revoltar contra injustiças, negociar coalizões, solucionar conflitos e cooperar ativamente. O próprio Frans de Waal apresenta essas capacidades como 'blocos de construção' da moralidade humana, progressivamente empilhados no curso dos últimos milhões de anos pela seleção natural. Contudo, essa concepção não incorreria na falácia naturalista, buscando *justificar comportamentos* humanos a partir de um *juízo de fato* sobre como eles evoluíram? Como o primatólogo lidaria com essa objeção significativa?

Frans de Waal enfrenta a objeção em artigo recente (2014). De acordo com ele, "ser" e "dever ser" são o yin e o yang da moralidade. Temos os dois, precisamos de ambos, não são os mesmos, mas, apesar disso, não podem ser completamente separados. Ambos se complementam (Waal, 2014, 186). De acordo com sua posição, devemos distinguir as perspectivas *conceituais* e *comportamentais* de exame da questão. Do ponto de vista conceitual, é realmente difícil abandonar uma distinção rígida entre os domínios do 'ser' e do 'dever ser', pois é injustificável derivar obrigações morais a partir de juízos sobre como as coisas 'são'. Contudo, quando as tendências comportamentais e motivações são postas em jogo, é difícil não reconhecer como a evolução, paulatinamente, imbuíu em certas espécies padrões de comportamento que podemos avaliar como normativos (Waal, 2014, 187).

O etólogo invoca, inclusive, a própria posição de David Hume para sustentar que a consideração de elementos biológico-evolutivos na explicação da moralidade não viola em si mesma a falácia naturalista. Refere, inclusive, o reconhecimento por parte do filósofo escocês de que a moralidade não é fruto apenas da racionalidade, mas também de emoções inatas que moldam nossa própria percepção das experiências sensoriais. A normatividade humana é construída sobre um arcabouço cognitivo evoluído e parcialmente compartilhado com outras espécies. De acordo com o argumento, a biologia, não pertence apenas ao domínio do 'ser', mas também investiga questões conectadas ao 'dever ser', na medida em que também estuda como diversas espécies de animais são capazes de se comportar com base em padrões *normativos* que fixam modelos (*templates*) a serem perseguidos. Tal conclusão pode ser extraída mesmo do comportamento de aranhas e cupins:

Que o comportamento animal não é livre de normatividade (definida como a adesão a um ideal ou padrão) dificilmente precisa de argumento. Tome a reação da aranha a uma teia danificada. Se o dano for extenso, ela abandonará sua teia, mas na maioria das vezes entrará em modo de reparo, trazendo a teia de volta ao seu estado funcional anterior, preenchendo buracos ou repondo fios danificados (...). Da mesma forma, perturbar um ninho de formigas ou cupins leva a reparos imediatos, assim como danos a uma represa de castor ou a um ninho de pássaros. A natureza está cheia de estruturas físicas construídas por animais guiados por um modelo de como a estrutura deve parecer. Este modelo motiva o reparo ou ajuste assim que a estrutura se desvia do ideal. Em outras palavras, os animais tratam essas estruturas de forma normativa (Waal, 2014, 187, tradução livre).

Frans de Waal reconhece que, inobstante tais comportamentos sejam observados, a questão de fundo atinente à falácia naturalista consiste em definir se outros animais são guiados por emoções similares ao sentimento de obrigação moral de observar certos valores que orienta o comportamento humano. Em sua perspectiva, uma miríade de outras espécies apresentam disposições de autocontrole comportamental a partir do entrelaçamento entre certas emoções e sua cognição. De acordo com Frans de Waal, a própria hierarquia de dominância observada entre os grandes primatas é um sistema de moralidade fundado nessa conexão e que se encontra, inclusive, entre os traços evolutivos que levaram à emergência da normatividade humana. Referindo como os chimpanzés se orientam a partir da hierarquia social, controlando seus impulsos por dominância em respeito à posição do macho alfa, o etólogo destaca como tanto na espécie humana quanto entre os chimpanzés, a mesma região do cérebro (o lobo frontal) é acionada para mediar a resposta comportamental de inibição do comportamento de modo a observar a hierarquia.

Além disso, outros comportamentos observados em chimpanzés também sinalizam a existência de capacidades cognitivas voltadas à cooperação. Como exemplos, menciona as disposições à reconciliação após uma situação de conflito; resolver preventiva de conflitos, com indicação de que a harmonia social é um valor observado no trato entre indivíduos; insurgência contra a desigualdade; e mesmo, embora com menos intensidade, à preocupação com a comunidade. De acordo com o autor, “há muitos exemplos de monitoramento imparcial e mediação que parecem refletir valores comunitários. Em algumas espécies, intervenções dos membros maior ranqueados no grupo encerram brigas ou ao menos reduzem a severidade da agressão” (Waal, 2014, 198, tradução livre). Essas evidências indicariam a prática de ações em benefício

da comunidade como um todo, inclusive lastreadas na reputação de cada indivíduo em suas relações sociais (Waal, 2014, 200).

Com esses fundamentos, de Waal sustenta a dificuldade de considerar as observações de biólogos como pertencentes exclusivamente ao domínio do 'ser'. Explicações sobre o comportamento dos animais sem atribuir intenções, objetivos ou mesmo valores morais são insuficientes por perderem aspectos fundamentais de suas experiências (Waal, 2014, 200). Trata-se de argumento similar ao de Herbert Hart em seu livro clássico, *The Concept of Law* (Hart, 1994[1961]), onde distingue entre dois pontos de vista teóricos sobre o direito. Do ponto de vista *externo*, como o adotado por teorias sociológicas, apenas comportamentos são observados, sem referência a qualquer estado mental ou motivo interno dos cidadãos e profissionais de uma comunidade jurídica. O ponto de vista *interno*, por sua vez, alude à atitude prática de quem respeita as regras, endossando um padrão convergente de comportamento. Assim, o ponto de vista interno refere-se a um tipo específico de atitude normativa mantida por aqueles que aceitam a legitimidade das regras. Examinar o direito apenas do ponto de vista externo implicaria perder uma riqueza de motivações que levam as pessoas a respeitar as normas jurídicas, endossando sua validade.

Assim como Hart, de Waal distingue entre os pontos de vistas interno e externo nas explicações etológicas. Relevar as disposições mentais dos animais – e, particularmente, de animais mais próximos dos humanos na escala evolutiva - na explicação do comportamento altruísta implicaria deixar de lado inúmeras nuances que apenas podem ser afirmadas a partir da premissa de que, como nós, eles são capazes de ter objetivos, valores e sentir emoções.

É preciso destacar que a abordagem de Frans de Waal não apresenta um importante argumento quanto ao obstáculo da falácia naturalista na ética evolucionista. A falácia está em empreender o salto lógico entre a *descrição* de um determinado estado de coisas e a *prescrição* de uma ação em particular. A explicação evolucionista da moralidade, contudo, não precisa incorrer nesse erro. É perfeitamente possível descrever *como* a capacidade cognitiva moral evoluiu sem *justificar* qualquer comportamento específico.

Uma tal explicação poderia apresentar, inclusive, as razões evolutivas que levam nossa cognição a preferir um curso de ação em detrimento de outros, ou até predisposições que estão na raiz de certos juízos morais. Considere, por exemplo, a alegação de que tendemos a preferir alimentos mais gordurosos em virtude de nosso passado ancestral, onde tais alimentos eram mais escassos e energeticamente mais substanciosos. Esta,



por si só, é a afirmação de um estado de coisas evolutivos, que não incorpora qualquer *prescrição* a respeito de nosso comportamento alimentar. A falácia estaria em concluir que, em razão de tal história evolutiva, devemos continuar a consumir alimentos gordurosos atualmente.

Além disso, invocar a história evolutiva é um passo relevante para compreendermos adequadamente como a própria cisão entre 'ser' e 'dever ser' emergiu na história evolutiva humana. Até a evolução da classe dos mamíferos, possivelmente nenhum outro animal era capaz, cognitivamente, de orientar suas ações com base em um conjunto de regras axiologicamente descritíveis. A escolha dos mamíferos como premissa é justificada com base nas evidências, já mencionadas, de que muitas espécies pertencentes a essa classe apresentam disposições intrincadas e necessárias para a cooperação.

A cognição moral - a que venho denominando de 'mente normativa' (Almeida, 2011) - pode ter emergido como um conjunto especializado de capacidades intelectuais funcionalmente diferenciadas e dedicadas a avaliar situações socialmente complexas e a decidir pelo curso de ação mais adaptativo. A exemplo de outros domínios intelectivos que se especializaram em nossa história evolutiva, como os voltados à compreensão das relações de causa/efeito no mundo físico, à visão, à audição ou, mais recentemente, à aquisição da linguagem, a cognição moral pode ter emergido como resultado da pressão social da vivência em comunidades sociais progressivamente maiores e mais complexas (Aiello & Dunbar, 1993; Dunbar, 1998).

Em uma perspectiva evolutiva, a cognição moral pode ter se originado de traços intelectivos ligados à compreensão do mundo físico, como os *modelos* a que alude Frans de Waal ao referir-se ao modo de reparo das teias das aranhas ou de represas de castores. Posteriormente, e progressivamente, se estabeleceu na arquitetura cerebral a distinção funcional entre a cognição do mundo físico e a cognição moral. Com o desenvolvimento histórico da lógica, as operações mentais do primeiro tipo de cognição resultam em juízos de 'ser', ao passo que operações relativas à segunda categoria, por sua vez, geram juízos de 'dever ser'.

Tal explicação não incorre na falácia naturalista, na medida em que apenas explica a cadeia de causalidade que explica a emergência da capacidade humana de formular juízos normativos, sem que daí resulte qualquer justificção de comportamentos específicos. Evidentemente, contudo, ao explicar a emergência da cognição moral com referência ao contexto adaptativo de sua evolução, essa abordagem estabelece um entrelaçamento indelével entre os juízos morais e a estrutura de nossa psicologia. Apenas somos capazes de formular juízos de 'dever ser' por

termos uma mente normativa, mas a *justificativa* de tais juízos deve ser debatida em seus próprios termos.

#### 4. CONCLUSÃO

Frans de Waal é referência fundamental para discutir os fundamentos do direito a partir de uma referência naturalista. Suas descobertas nos campos da etologia e, mais especificamente, da primatologia, têm inegavelmente avançado a compreensão sobre o comportamento animal e denotando inúmeros aspectos de similitude com o nosso próprio comportamento.

O reconhecimento de que os membros de outras espécies são capazes de sentir gratidão, empatia e compaixão em relação ao sofrimento de outros, reagir ao tratamento injusto, compor e mediar conflitos, e agir em relações recíprocas, tem o potencial de revolucionar o modo como compreendemos o direito. Tradicionalmente, a teoria jurídica – assim como a maior parte da filosofia moral – se baseia em pressupostos antropocêntricos, que consideram a moralidade e o direito como fruto da racionalidade humana e independente de nossa natureza animal.

O presente artigo objetivou apresentar algumas das contribuições do professor holandês para a teoria jurídica. A dificuldade de enfrentamento do tema, dada a diversidade de públicos a quem a discussão pode interessar, é evidente. De um lado, os juristas não estão acostumados ao enfrentamento da literatura em etologia, primatologia ou antropologia evolutiva, o que explica em parte a profunda desatualização científica dos livros-texto na área jurídica. Essa circunstância é compreensível, diante da especialização em torno de temas relacionados às preocupações mais rotineiras de professores, pesquisadores e profissionais da área. Por outro lado, filósofos e cientistas também desconhecem em grande parte as construções teóricas dos juristas sobre temas que podem ser informados com base na literatura científica. Tal cenário tem levado à ausência de um diálogo interdisciplinar importante, que pode levar à revisitação de temas mais fundamentais da teoria jurídica.

Com vistas a construir pontes entre áreas tão distintas, o artigo buscou expor, em linguagem tão acessível quanto possível a ambos os públicos, alguns dos caminhos teóricos que a obra de Frans de Waal abre para a teoria do direito. Sem qualquer pretensão de esgotar o tema dada a proficuidade de temas investigados por Frans de Waal, a primeira seção destinou-se a expor pontos essenciais das pesquisas do primatólogo, apresentando evidências importantes sobre a vida moral dos animais.

Essa exposição, por si só, já denota a contribuição do cientista para a teoria jurídica, na medida em que revela o caráter antropocêntrico das discussões da doutrina tradicional em direito, que usualmente assume, sem qualquer investigação teórica mais profunda, a normatividade como um traço exclusivo dos membros da espécie humana. Como denota a obra de Frans de Waal, trata-se de premissa que demanda um trabalho de justificação e de construção de distinções entre nós e outros animais. Talvez seja necessário, inclusive, categorizar diferenças entre os outros animais, como os grandes primatas, golfinhos e elefantes, entre tantas outras espécies que apresentam nível intelectual elevado e, por isso mesmo, demandam proteção específica de sua singularidade a fim de que possam florescer de modo adequado.

A segunda seção do texto, por sua vez, pretendeu apresentar três pontos de impacto da obra de Frans de Waal à teoria jurídica: as contribuições de suas investigações para a compreensão da cognição moral como fundamento do direito; os direitos dos animais; e a superação da falácia naturalista por teorizações naturalizadas sobre o espectro do direito. Evidentemente, outras contribuições poderiam ser sinalizadas, como a abertura para questionar noções basilares como as de sujeito de direito e de indivíduo, o próprio conceito de normatividade, ou, ainda, a própria concepção de razão prática como atributo exclusivo dos humanos. Contudo, espera-se, como contribuição do artigo, ter-se apresentado, ainda que brevemente, como a obra de Frans de Waal pode ser utilizada como premissa a discussões relevantes à teoria jurídica.

As contribuições de Frans de Waal são relevantes na medida em que indicam a necessidade de maior articulação entre teoria jurídica e estudos sobre comportamento animal. Seu trabalho pode ser considerado um exemplo de como tal articulação pode ser efetivamente concretizada, sobretudo quando se trata de compreender o papel da cognição moral na construção do direito. Desse modo, espera-se que a presente exposição possa servir como ponto de partida para outras reflexões sobre a relação histórica entre direito e comportamento animal, discutindo as contribuições, as limitações e as possíveis aplicações práticas que tais estudos possam oferecer para a compreensão e a aplicação do direito.

## REFERÊNCIAS BIBLIOGRÁFICAS

Aiello, Leslie C., Dunbar, Robin (1993). Neocortex size, group size, and the evolution of language. *Current Anthropology*, 34(2), 184-193. <https://doi.org/10.1086/204160>

- Almeida, Fábio Portela Lopes De (2011). *A evolução da mente normativa: origens da cooperação humana*. Brasília: Universidade de Brasília.
- Almeida, Fábio Portela Lopes De (2013). As origens evolutivas da cooperação humana e suas implicações para a teoria do direito. *Revista Direito GV*, 9(1), 243-268. <https://doi.org/10.1590/S1808-24322013000100009>
- Almeida, Fábio Portela Lopes De (2020). *Constitution: the Darwinian evolution of a societal structure*. Baden-Baden: Nomos. <https://doi.org/10.5771/9783748905547>
- Alvim, José Eduardo Carreira (2001). *Elementos de teoria geral do processo*. Rio de Janeiro: Forense.
- Brosnan, Sarah F. & De Waal, Frans (2003). Monkeys reject unequal pay. *Nature*, 425(6955), 297-299. <https://doi.org/10.1038/nature01963>
- Brosnan, Sarah F., Houser, Daniel, Leimgruber, Kristin, Xiao, Erte, Chen, Tianwen & de Waal, Frans, B. M. (2010a). Competing demands of prosociality and equity in monkeys. *Evolution and Human Behavior*, 31(4), 279-288.
- Brosnan, Sarah F., Talbot, Catherine, Ahlgren, Megan, Lambeth, Susan P. & Schapiro, Steven J. (2010b). Mechanisms underlying responses to inequitable outcomes in chimpanzees, Pan troglodytes. *Animal Behaviour*, 79(6), 1229-1237.
- Burghardt, Gordon M. (1985). Animal awareness: Current perceptions and historical perspective. *American psychologist*, 40(8), 905.
- Byrne, Richard W. & Whiten, Andrew (1988). *Machiavellian intelligence*. Oxford: Oxford University Press.
- Carter, Gerald (2014). The reciprocity controversy. *Animal Behavior and Cognition*, 1(3), 368-386. <https://doi.org/10.12966/abc.08.11.2014>
- Chediak, Karla (2006). O problema da falácia naturalista para o projeto de uma ética evolucionista. *Kriterion: Revista de Filosofia*, 47(113), 147-157. <https://doi.org/10.1590/S0100-512X2006000100008>
- Chernillo, Daniel (2013). *The natural law foundations of modern Social Theory*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511842511>
- Contreras, Francisco José (Org.) (2013). *The threads of natural law: unravelling a philosophical tradition*. London: Springer. <https://doi.org/10.1007/978-94-007-5656-4>
- Das, Marjolijn (2000). Conflict management via third parties: post-conflict affiliation of the aggressor. In Frans de Waal, & Filippo Aureli (Org.), *Natural conflict resolution* (pp. 263-280). Berkeley & Los Angeles: University of California Press. <https://doi.org/10.1525/9780520924932-019>

- Dunbar, Robin (1998). *Grooming, gossip, and the evolution of language*. Cambridge: Harvard University Press.
- Ferraz Junior, Tércio Sampaio (2013). *Introdução ao estudo do direito: técnica, decisão, dominação*. São Paulo: Atlas.
- Hart, Herbert L. (1994). *The concept of law*. Oxford: Clarendon Press.
- Hayek, F. A. (1998). *Law, Legislation and Liberty*. London: Routledge.
- Holmes, Jr Oliver Wendell (1900). Law in science and science in law. *The Brief*, 2, 105-127.
- Hume, David (2009). *Tratado da natureza humana*. São Paulo: UNESP.
- Jordan, Fiona M et al. (2013). Cultural evolution of the structure of human groups. In Peter Richerson & Morten Christiansen (Org.), *Cultural evolution: society, technology, language, and religion* (pp. 87-116). Cambridge: The MIT Press.
- Lorini, Giuseppe (2018). Animal norms: an investigation of normativity in the non-human social world. *Law, culture and the humanities*, 18(3), 652-673. <https://doi.org/10.1177/1743872118800008>
- Maine, Henry Sumner (1906). *Ancient law*. New York: Henry Bolt and Company.
- Moore, G. E. (1993). *Principia ethica*. Cambridge: Cambridge University Press.
- Nowak, Martin A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560-1563. <https://doi.org/10.1126/science.1133755>
- Nuccetelli, Susana (2021). *The Routledge guidebook to Moore's Principia Ethica*. London: Routledge. <https://doi.org/10.4324/9780429275975>
- Plotnik, Joshua M. et al. (2011). Elephants know when they need a helping trunk in a cooperative task. *Proceedings of the National Academy of Sciences*, 108(12), 5116-5121. <https://doi.org/10.1073/pnas.1101765108>
- Plotnik, Joshua M., Waal, Frans B. M. De, & Reiss, Diana (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences*, 103(45), 17053-17057. <https://doi.org/10.1073/pnas.0608062103>
- Posner, Richard A. (1980). A Theory of primitive society, with special reference to law. *The Journal of Law and Economics*, 23(1), 53. <https://doi.org/10.1086/466951>
- Preston, Stephanie D.; WAAL, Frans De. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, v. 25, n. 1, p. 1–20, 2001. <https://doi.org/10.1017/S0140525X02000018>
- Richardson, Peter J. & Boyd, Robert (2008). *Not by genes alone*. Chicago: University of Chicago Press.

- Sacco, Rodolfo (1995). Mute law. *The American Journal of Comparative Law*, 43, 455-467. <https://doi.org/10.2307/840648>
- Santos, Moacyr Amaral (2011). *Primeiras linhas de direito processual civil* (vol. 1). São Paulo: Saraiva.
- Silk, Joan B. (2009). Nepotistic cooperation in non-human primate groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3243-3254. <https://doi.org/10.1098/rstb.2009.0118>
- Singer, Peter (2009). *Animal liberation*. New York: Open Road.
- Soltis, Joseph, Boyd, Robert, & Richardson, Peter J. (1995). Can group-functional behaviors evolve by cultural group selection? An empirical test. *Current Anthropology*, 36(3), 473-494. <https://doi.org/10.1086/204381>
- Spada, Emanuela Cenami (1997). Amorphism, mechanomorphism, and anthropomorphism. In Mitchell, R. W., Thompson, N. S. and Lyn Miles, H. (Eds.), *Anthropomorphism, anecdotes and animals* (pp. 37-49). Albany, NY: State University of New York Press.
- Waal, Frans De (1989). *Chimpanzee politics: power and sex among apes*. Baltimore: Johns Hopkins Paperbacks.
- Waal, Frans De (1991). The Chimpanzee's sense of social regularity and its relation to the human sense of justice. *American Behavioral Scientist*, 34(3), 335-349. <https://doi.org/10.1177/0002764291034003005>
- Waal, Frans De (2000). Primates: a natural heritage of conflict resolution. *Science*, 289, 586-590. <https://doi.org/10.1126/science.289.5479.586>
- Waal, Frans De (2009). Darwin's last laugh. *Nature*, 460(7252), 175-175.
- Waal, Frans De (2010). *Age of empathy: nature's lessons for a kinder society*. New York: Harmony Books.
- Waal, Frans De (2014). Natural normativity: the 'is' and 'ought' of animal behavior. *Behaviour*, 151(2-3), 185-204. <https://doi.org/10.1163/1568539X-00003146>
- Waal, Frans De et al. (2006). *Primates and philosophers: how morality evolved*. Princeton: Princeton University Press.
- Waal, Frans De, van Hoof, Jan A. R. A. M., & Netto, Willem J. (1976). An ethological analysis of types of agonistic interaction in a captive group of Java-monkeys (*Macaca fascicularis*). *Primates*, 17(3), 257-290. <https://doi.org/10.1007/BF02382785>
- Wright, Robert (2006). The uses of anthropomorphism. In Frans De Waal et al., *Primates and philosophers: how morality evolved*. Princeton: Princeton University Press. <https://doi.org/10.1515/9781400830336-007>

eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.31500>

# MECANISMOS SOCIALES, ANTROPOMORFISMO Y PROCESOS COGNITIVOS EN ANIMALES NO HUMANOS. A 40 AÑOS DE *LA POLÍTICA DE LOS CHIMPANCÉS*

## *Social Mechanisms, Antropomorphism and Cognitive Processes in Non-Human Animals. 40 Years after Chimpanzee Politics*

Oscar David CAICEDO MACHACON   
*Universidad del Atlántico, Colombia*  
[oscarcaicedo@mail.uniatlantico.edu.co](mailto:oscarcaicedo@mail.uniatlantico.edu.co)

Rene J. CAMPIS   
*Universidad del Atlántico, Colombia*  
[renecampis@mail.uniatlantico.edu.co](mailto:renecampis@mail.uniatlantico.edu.co)

Eduardo BERMÚDEZ BARRERA  
*Grupo de Investigación Holosapiens, Universidad del Atlántico, Colombia*  
[eduardobermudez@mail.uniatlantico.edu.co](mailto:eduardobermudez@mail.uniatlantico.edu.co)

Recibido: 29/06/2023    Revisado: 30/07/2023    Aceptado: 12/08/2023

**RESUMEN:** La atribución de pensamientos y mente a animales no humanos sigue generando debates y controversias entre los estudiosos del comportamiento animal. Sumado a ello, los mecanismos sociales -el *juego* como un comportamiento producto de la selección natural con valor de supervivencia- llaman poderosamente la atención de etólogos cognitivos,

psicólogos y filósofos de la mente a la hora de establecer hasta dónde nos es permitido abrazar el antropomorfismo científico sin caer en analogías apresuradas del antropomorfismo ingenuo y desinformado.

Este artículo aborda el problema de cómo estudiar algunas estrategias de supervivencia en grupos sociales, cómo justificar el antropomorfismo científico y hasta qué punto podemos hablar de procesos mentales en otros animales, apelando al concepto de *mecanismos sociales*, entre otros, en de Waal.

*Palabras clave:* inteligencia social, teoría de la mente, etología cognitiva, etología social.

**ABSTRACT:** The attribution of thoughts and mind to non-human animals generates heated debates and controversies among scholars of animal behavior. Social mechanisms –such as play as a behavior produced by natural selection that has survival value– strongly catch the attention of ethologists, psychologists and philosophers of mind when it comes to establishing how far can we go in embracing scientific anthropomorphism without falling into hasty analogies of naive and uninformed anthropomorphism.

This paper appeals to the concept of *social mechanisms*, among others, in de Waal to address the problem of how to study some survival strategies in social groups, how to justify scientific anthropomorphism and to what extent we can talk about mental processes in other animals.

*Keywords:* social intelligence, theory of mind, cognitive ethology, social ethology.

## 1. INTRODUCCIÓN

Un presupuesto básico de *La política de los chimpancés* es la vida en sociedad. Hablar de *política* en primates o cualquier otra especie animal solo cobra sentido cuando se entiende que nos estamos refiriendo a animales sociales. La sociabilidad de variadas especies animales ha sido trabajada en extenso por múltiples etólogos, psicólogos y biólogos, y Frans de Waal es, sin duda, una de las referencias más citadas.

En la segunda mitad del siglo XX, aparece la expresión *Etología social* en Crook y Goss-Custard (1972) para referirse al estudio del comportamiento social de diversas especies e individuos que viven en grupos permanentes, en relación con el medio ecológico. Colmenares (2014) plantea que una definición de comportamiento social neutral desde el



punto de vista teórico tendría que considerar que un comportamiento es social siempre que tenga como receptor a otro individuo de la misma especie. Esta definición excluye comportamientos autodirigidos, dirigidos hacia elementos del medio, o hacia individuos de otras especies, siempre y cuando tales acciones y sus consecuencias no proporcionen claves informativas inmediatas o diferidas a otros individuos de la especie. Una definición más amplia podría plantearse en términos de que un comportamiento es social en tanto se pueda constatar que afecta y potencialmente modifica la conducta (o el estado fisiológico o psicológico) de otro individuo de la especie. Esta es una definición eminentemente funcional.

Hay, sin embargo, una tercera definición que enfatiza el sustrato cognitivo de la causa proximal, incorporando variables inobservables como la intención o la atribución de conocimiento por parte del emisor, aunque tal inferencia se realiza a partir del análisis de observables. Aquí, el emisor de un comportamiento social actúa de un modo u otro animado por la intención de estimular al receptor.

¿Cómo podríamos abordar, precisamente, el tema de la etología social desde la obra de Frans de Waal?

## 2. EL JUEGO COMO MECANISMO SOCIAL

El último capítulo de su libro *La política de los chimpancés*, Frans de Waal lo titula, sugerentemente, “Los mecanismos sociales”. Ahí, el autor revisa algunos de estos mecanismos, introduciendo el concepto de *conciencia triádica*, entendida, según él, como un requisito básico de toda estructura jerárquica basada en coaliciones –además de reconocer a los demás miembros de un grupo individualmente–. Mientras que el reconocimiento individual garantiza la existencia de una jerarquía bien establecida (en la que cada individuo conoce el lugar que le corresponde en su grupo), la conciencia triádica es la capacidad de percibir las relaciones sociales que se dan entre otros individuos y formar relaciones triangulares, no sólo del tipo “yo me puedo relacionar con B y con C (por separado)”, sino también de saber que B y C tienen una relación, y saber que puedo sacarle provecho (de Waal, 1989).

Pues bien, el juego es uno de los mecanismos sociales por medio del cual humanos y otros animales sociales se inician en el proceso de reconocimiento de jerarquías y el refinamiento de la conciencia triádica, además de, como es de suponer, el respeto de reglas y normas establecidas.

Los comportamientos son considerados por la etología como *adaptaciones biológicas efectivas* en el contexto del animal que las realiza. Comprender tales adaptaciones no es tarea fácil, máxime cuando un mismo comportamiento puede ser útil para diversos propósitos en diferentes situaciones. Algunos autores como Asensio (2014) consideran que el juego o comportamiento lúdico presenta claramente esta dificultad. Al ser característico de la mayoría de los mamíferos y aves, no resulta difícil pensar en el juego como un comportamiento producto de la selección natural con un evidente valor de supervivencia y, por tanto, beneficioso para quienes lo practican, comportamiento que posee algunas características estructurales (actos, señas, duración, intensidad), causales (contextos, situaciones, estímulos) y funcionales (sentido adaptativo).

Si hay algo que caracteriza a los mamíferos probablemente sea la propensión al juego y, si hay algo que nos caracteriza a los humanos, posiblemente sea nuestra propensión al juego aún en etapa adulta y el establecimiento de reglas previas comparables, en muchos casos, a un sistema de castigo (cuando se transgreden las normas) /recompensa (cuando se acierta). En otras especies de primates como los chimpancés, por ejemplo, en crías más pequeñas, el aprendizaje y el juego son de vital importancia para la adquisición de habilidades en el uso de herramientas. Cuando en cautiverio se les niega a las crías el acceso a palos o piedras para jugar, su capacidad para resolver problemas utilizando tales instrumentos en la edad adulta, se ve considerablemente reducida (Llorente, 2019).

Diversas especies han ido moldeando y manteniendo el comportamiento lúdico a lo largo de la evolución, pero, ¿cuál es la razón? Jugar implica un coste biológico por su enorme inversión energética, además de exponer a los individuos al riesgo de la depredación, heridas, o incluso, la muerte. Es de suponer que los costes deben ser superados por los beneficios; de lo contrario, sería una desventaja, y habría desaparecido a lo largo de la evolución. Parece cada vez más claro que el juego no solamente es terapéutico –sustituyendo en ocasiones ansiedad y angustia por felicidad y alegría–, sino que también nos hace flexibles cognitivamente, preparándonos para afrontar cambios en el entorno, y haciéndonos animales más inteligentes. La conducta lúdica podría ser la base de comportamientos sociales más complejos como la cooperación, e incluso el germen para capacidades como la moralidad y la empatía (Llorente, 2019). En diversas especies de primates –nosotros incluidos–, el juego fomentaría la aparición de conductas innovadoras y creativas, facilitando el surgimiento de comportamientos y tradiciones culturales, teniendo en cuenta que una de las funciones biológicas principales del juego reside en la adquisición de habilidades y experiencias útiles para la edad adulta.

El juego permite a los individuos comprometerse y vincularse con su propio entorno, aprendiendo a interactuar con él y a aprender de él, reconociendo objetos y comprendiendo para qué sirven, descubriendo que a veces hay cosas que están escondidas detrás de otras, y aprendiendo qué se puede hacer y qué no con otros individuos de mi grupo [...] Otras de las adquisiciones importantes que se consiguen con el juego tienen que ver con la adquisición de capacidades cognitivas y de resolución de problemas, la consolidación de las relaciones sociales o el ajuste y puesta a punto de la musculatura y del sistema nervioso a través del entrenamiento físico (Llorente, 2019, 292)

Aunque *La política de los chimpancés* no ofrece un concepto de juego, de Waal ilustra algunas conductas lúdicas entre los chimpancés de su colonia, indagando precisamente sobre lo fundamental del juego en otros primates para entender parte de nuestro propio comportamiento. Sugiere que cuando las relaciones de dominancia existentes en el grupo se burlan durante el juego, no hay peligro de que surjan confusiones, puesto que ya se han dejado suficientemente claras en otros momentos. No pone en riesgo la jerarquía del grupo el hecho de que, durante el juego, una cría de chimpancé "someta" durante una pelea ficticia a un macho de mayor rango, en lo que siempre parece indispensable tener "cara de juego" (*play face*). "La cara de juego, la mueca con la boca abierta hacia los lados y el gesto de pedir con la mano no son imitaciones de conductas humanas, sino formas naturales de comunicación no verbal que los chimpancés y los humanos tenemos en común" (de Waal, 1989, 33). La cara de juego es útil además para dejar claro que se está jugando.

Un ejemplo podría ilustrar claramente lo que, en párrafos anteriores, mencionamos como *conciencia triádica*. Afirma de Waal (1989) que, en la colonia de Arnhem, dos madres, Jimmie y Tepel, están sentadas mientras sus dos hijos juegan a sus pies con la arena (poniendo caras de juego, luchando y tirándose arena). Entre las dos madres se encuentra también la hembra de mayor edad del grupo, Mamá, quien duerme. De un momento a otro, los pequeños empiezan a pelear, golpeándose y tirándose del pelo. Jimmie les regaña con un suave gruñido de amenaza, y Tepel, con notable intranquilidad, cambia de posición. Las dos crías siguen peleándose y, finalmente, Tepel despierta a Mamá tocándole las costillas con el dedo. Tan pronto como Mamá despierta, Tepel le muestra a las dos crías que siguen peleando. Cuando Mamá da un paso hacia adelante en actitud de amenaza, agitando un brazo en el aire con ruidosos ladridos, las crías dejan de pelearse. Mamá se recuesta nuevamente y continúa durmiendo.

De Waal sugiere atender al menos dos cosas en este ejemplo. Por un lado, que Mamá es la hembra de mayor rango de la colonia y, por tanto,

muy respetada; y segundo, que generalmente los conflictos entre crías engendran tanta tensión entre las respectivas madres, que al final ellas también pueden estallar y empezar a pelearse. La causa de esta tensión es probablemente el hecho de que cada madre intenta evitar que la otra intervenga en la pelea de las crías. En este caso, cuando el juego de los pequeños se convirtió en pelea, ambas madres se encontraron en una situación difícil. Tepel resolvió el problema activando una tercera parte dominante, Mamá, y señalándole el problema. Obviamente, Mamá notó de inmediato que se esperaba que ella interviniese.

Pero de Waal va más allá, extrapolando situaciones similares a la escena política, introduciendo el término *inteligencia estratégica*. Cuando un país tiene tensiones con otro menos fuerte, el tercero en discordia hace generalmente proposiciones al que se encuentra en situación más débil. Ha habido muchos procesos como éstos en la historia de la civilización, que han ayudado a mantener el equilibrio de poder otorgando a la tercera potencia una posición influyente. Si un país no quiere quedarse al margen, debe colocar su peso en el lado más ligero de la balanza.

El mismo principio se aplica a la psicología social y se conoce con el nombre de "formación de coaliciones con ventaja mínima". Se ha comprobado mediante experimentos que, cuando en un juego en el que participan tres jugadores se da al más débil la oportunidad de ganar puntos a su favor aliándose o con el jugador más fuerte o con aquel que ocupa una posición intermedia, preferirá aliarse con este último (de Waal, 1989, 228).

Hace ya varios años que la psicología social explota estos temas. Las teorizaciones sobre la naturaleza humana, las estrategias grupales e individuales para sacar ventaja en situaciones específicas, las explicaciones sobre el egoísmo y el altruismo y las diferentes variaciones sobre el dilema del prisionero, son algunos ejemplos. Sin embargo, está claro que extrapolar tales explicaciones a otras especies de animales sociales divide opiniones. Las acusaciones de antropomorfismo están presentes en los debates al respecto.

### **3. ANTROPOMORFISMO, CANON DE MORGAN (Y DE GLOCK) Y EVIDENCIA ANECDÓTICA**

El antropomorfismo y su defensa –aunque matizada– es tema recurrente en la obra de de Waal. El antropomorfismo se ha relacionado con la evidencia anecdótica del siglo XIX para explicar la continuidad entre seres humanos y otros animales. Ha habido, sin embargo, un resurgimiento

reciente del interés en el antropomorfismo, atribuible en parte a dos acontecimientos: el aumento de estudios en (i) etología cognitiva, y (ii) los requisitos de las diversas formas para la aplicación de una ética ambiental (derechos de los animales, etc.). El primero, la investigación sobre la vida mental y el comportamiento de los animales, resurgió, desde la década del 70 del siglo pasado, con un nuevo énfasis en las experiencias mentales de los animales, sobre todo en su entorno natural, en el curso de su vida cotidiana. El paso hacia la atribución de estados mentales a animales no humanos ha sido polémico desde sus inicios. ¿Qué significa decir que un animal tiene acceso a sus propios pensamientos, y cómo se podría establecer experimentalmente a la verdad o falsedad de tal afirmación? (Mitchell, 2005). Una versión fuerte de antropomorfismo se encuentra en algunos defensores de la etología cognitiva que intentan explicar los comportamientos de los animales no humanos apelando a estados mentales similares a aquellos con los que explicamos nuestro propio comportamiento. Además, están los “excepcionalistas humanos” y antropocentristas, como les llama Glock, que se oponen a la idea del pensamiento animal o intentan trazar distinciones cualitativas entre humanos y otros animales.

La continuidad evolutiva no puede hacernos desconocer diferencias que han surgido en el camino de la evolución. De la continuidad biológica no se puede inferir que los demás animales deban necesariamente estar próximos a nuestras experiencias mentales. Pequeñas diferencias bioquímicas en el genotipo pueden dar lugar a grandes diferencias en el fenotipo.

Causey y Bjorklund (2016), entre otros, sugieren que las principales diferencias entre humanos y demás animales están relacionadas principalmente con la habilidad de usar el lenguaje, de heredar cultura material con precisión y de razonar científicamente. Mientras el resto de animales se adaptan para encarar los retos del ambiente, los humanos han logrado –para bien o para mal– adaptar el mundo para satisfacer sus necesidades. Si bien no parece haber mucho desacuerdo sobre la continuidad entre otros animales y los humanos en muchos aspectos, las discusiones científicas y filosóficas sobre la continuidad en los rasgos cognitivos están lejos del consenso.

La aproximación biológica al estudio de la inteligencia se caracteriza básicamente por adoptar dos principios generales. El estudio de la inteligencia desde una perspectiva biológico-evolutiva tiene un interés fundamental en comparar la inteligencia entre distintas especies animales; el método comparativo parece de nuevo fundamental. En segundo lugar, se toma como premisa que la inteligencia está organizada por módulos, en

contraposición a una inteligencia más global o general. En este orden de ideas, las diferentes especies animales poseerían distintas inteligencias – diferentes módulos– en función de la presión recibida en los distintos nichos ecológicos y en función del desarrollo de diferentes estrategias de adaptación (Llorente, 2019).

Uno de los ejemplos que ofrecen los autores sobre la ruptura cognitiva entre humanos y otras especies es la autoconciencia. Aunque no parece problemático aceptar que algunas aves y mamíferos tengan un cierto grado de conciencia (recuérdese el test del espejo), es más complejo asegurar que alguna especie –a excepción de la humana– tenga autoconciencia en sentido pleno. Parece claro que los animales no son capaces de reflexionar sobre su propia condición de sujetos (sometidos a una ineludible temporalidad y caducidad) o sobre las particularidades más sutiles de sus procesos mentales, aunque sí que pueden tener un conocimiento rudimentario de su propia y peculiar individualidad, al menos en el aspecto corporal. Nos parece difícil imaginar que un perro, un gato e incluso un cuervo o un chimpancé *reflexionen* sobre su propia existencia y sus propios pensamientos.

Debido a que los humanos tienen la capacidad de reflexionar sobre experiencias pasadas, imaginar escenarios futuros y examinar necesidades y deseos actuales y futuros, pueden controlar los impulsos que resultan de sus propias representaciones del mundo y analizar las consecuencias a largo plazo de sus decisiones. De esta forma, los humanos cambiaron la naturaleza de las presiones selectivas al convertirse ellos mismos en una; así, desarrollaron su cognición social. Ésta es una capacidad única del *Homo sapiens* (Causey y Bjorklund, 2016, 42-43).

Sin embargo, la ausencia de evidencia no es evidencia de ausencia. No contar actualmente con la certeza de que otras especies animales tengan autoconciencia –en los términos anteriormente planteados– puede deberse a múltiples factores, como la barrera interespecie, la imposibilidad de otros animales de verbalizar sus estados mentales (en el caso de que los posean), no disponer de herramientas tecnológicas que permitan indagar empíricamente sobre tales estados mentales, o un posible error humano que vicie los experimentos conducentes a obtener tales evidencias.

¿Podemos cruzar la barrera interespecie y hablar, por ejemplo, de pensamiento en animales no humanos?, ¿pueden los animales no humanos tener pensamientos sin lenguaje? y, más aún, ¿podemos hablar de pensamiento conceptual en animales sin caer en el antropomorfismo extremo?

Hasta ahora, las respuestas a estas preguntas no son concluyentes. Generalmente, tales respuestas se alternan con acusaciones de

antropomorfismo, antropocentrismo y antroponegación. Mientras que el primero sugiere la posibilidad de atribuir capacidades psicológicas, sociales o normativas humanas a animales no humanos, el antropocentrismo basa las explicaciones de tales capacidades en los animales tomando como medida o referencia de las mismas las capacidades cognitivas humanas. La antroponegación puede entenderse como el rechazo absoluto al antropomorfismo, incluso en sus variantes más débiles. Entre los partidarios de la *antroponegación*, están a su vez los escépticos, particularmente preocupados por el antropomorfismo: los escépticos categóricos (que piensan que la investigación sobre la cognición animal no puede ser buena ciencia), y los escépticos selectivos (que piensan que algunos tipos de atribuciones no están realmente justificadas) (Andrews y Huss, 2014).

C. Lloyd Morgan estableció en lo que hoy conocemos como el *Canon de Morgan* que una acción no se debe interpretar como resultado del ejercicio de una facultad psíquica superior si puede interpretarse como el resultado del ejercicio de una actividad psíquica inferior en la escala psicológica (Allen-Hermanson, 2005).

No atribuir procesos mentales complejos a los demás animales cada vez que fuese posible explicar su comportamiento en términos más simples, y tener cuidado en las inferencias hechas a partir de la conducta animal (tratando siempre de hallar una explicación “más simple” al fenómeno observado) ofrecen prevenciones a la hora de establecer hasta qué punto un comportamiento es resultado de un proceso mental complejo. No atribuir “demasiada inteligencia” a otros animales, en tanto la explicación de un comportamiento dado pueda entenderse, por ejemplo, en términos de instinto y de asociación.

Sober (2005) sugiere que para entender el Canon en su contexto histórico, debemos entender contra qué estaba reaccionando Morgan. Darwin había argumentado a favor de la continuidad mental de humanos y organismos no humanos [“No hay ninguna diferencia fundamental entre el hombre y los animales superiores en sus facultades mentales. [...] Los animales inferiores, como el hombre, manifiestamente sienten placer y dolor, la felicidad y la miseria” (Darwin, 1936[1871], 448)]. Su sucesor, George Romanes, también hizo hincapié en esta idea. El objetivo de Darwin fue mostrar que las ideas evolucionistas aplican a características mentales no menos que a rasgos morfológicos y fisiológicos. Si todos los seres vivos están relacionados genealógicamente, podemos ubicar la aparición de novedades en las ramas interiores de los árboles filogenéticos en el que las puntas representan especies actuales y los nodos interiores representan antepasados comunes. Dada la ascendencia común

y el gradualismo, que era parte de la concepción de Darwin de la selección natural, las especies contemporáneas deben exhibir similitudes. Por tanto, llevar este punto de vista evolutivo hacia fenómenos mentales significa que debe haber continuidades psicológicas entre los humanos y otros animales.

Darwin no solo subordina al ser humano como especie biológica a la evolución de los organismos en *The Descent of Man* en lo referente a rasgos morfológicos, sino que defiende la evolución filogenética de las potencialidades psíquicas o espirituales del *Homo sapiens*. Las cualidades y atributos espirituales (intelectuales) de los humanos como el lenguaje, la cultura y la moral, son resultado de la evolución. Se propone describir y comparar las formas del comportamiento animal y salvar el abismo entre humanos y otros animales. Las vivencias mentales no están circunscritas al animal humano; están sometidas, como cualquier otra función biológica, a la evolución (Wuketits, 1984).

Darwin, citando la continuidad evolutiva de todas las especies existentes, consideró apropiado utilizar un lenguaje descriptivo continuo, ya sea hablando de seres humanos, perros o pulpos. Por tanto, en su visión, bien podría ser científicamente apropiado hablar de juego entre felinos, cuervos o cánidos.

Hay dos aspectos de las ideas de Darwin sobre la continuidad mental que merecen mención. El primero es un compromiso de continuidad mental sobre la historia evolutiva, basado en un compromiso con la idea de modificación por descendencia y selección de los fenotipos físicos y de comportamiento. El segundo es el compromiso con la continuidad mental entre organismos actuales, basado en la atribución de estados mentales por el método de la observación. Las ideas generales de Darwin sobre la continuidad entre especies no fueron bien recibidas por sus contemporáneos. Sus afirmaciones acerca de la continuidad mental fueron especialmente controvertidas, ya que se tomaron para socavar la idea de que los seres humanos son los únicos seres racionales y, por tanto, para socavar la separación moral entre los humanos y demás animales (Allen y Bekoff, 1999).

Al notar que los científicos a menudo subestimaban las capacidades mentales de los animales no humanos, usó libremente anécdotas, atribuyendo estados cognitivos a muchos animales sobre la base de observaciones de casos particulares en vez de experimentos controlados; es lo que algunos han dado en llamar "cognitivism anecdótico" (Jamieson y Bekoff, 1993).

Como Darwin, Morgan también quiso defender la hipótesis evolutiva de que toda la vida está genealógicamente relacionada, pero vio que no



se requiere obviar las diferencias que separan a humanos de otros animales. Una rama de un árbol filogenético puede desarrollar novedades que no surgen en otros; una genealogía común no requiere ausencia de diferencias cualitativas entre rasgos exhibidos por especies relacionadas. Morgan sostuvo que atribuir estados mentales a los demás depende de un examen introspectivo de uno mismo. Cuando llevo una taza a los labios, esto es porque creo que la taza contiene un líquido que deseo beber. Cuando veo a otro ser humano realizar la misma acción, infiero una causa mental similar. Morgan vio que este patrón de inferencia se extiende a través de los límites entre especies. Lo que Morgan llamó el “método inductivo doble” nos lleva a interpretar el comportamiento de los organismos en otras especies como derivado de las mismas causas que mueven a los seres humanos a la acción. Sin embargo, Morgan no concluyó que esta inducción justificara el antropomorfismo; por el contrario, consideró que daba lugar a un sesgo antropomórfico que requería un contrapeso: ese fue el papel que jugó el canon en su pensamiento (Sober, 2005).

El temor a que la propensión humana a antropomorfizar pudiera llevarnos a aceptar atribuciones humanas a otros animales donde nada semejante al fenómeno humano ocurre fue, si se quiere, el origen del Canon. La primera pista falsa que tenemos que identificar y descartar es el *anecdotalismo desinformado*, como se ha comentado de Darwin o en el que incurrió Konrad Lorenz.

Aunque los científicos que estudian el comportamiento animal habitualmente hacen uso de anécdotas y antropomorfismo en su vida cotidiana y en los libros para el público en general, a menudo explícitamente repudian la evidencia anecdótica y la interpretación antropomórfica publicada en el discurso científico. Esta ambivalencia entre los científicos sugiere que anecdotalismo y antropomorfismo son importantes para pensar o entender el comportamiento animal. Sin embargo, si se quiere que la investigación se considere como científica, deben ser eliminados. Algunos como Rollin (1997) sostienen, que eliminar el anecdotalismo y el antropomorfismo es innecesario. Ambos, a pesar de sus malas reputaciones, no son significativamente diferentes.

Los modelos antropomórficos muestran similitudes entre humanos y no humanos. Deben ser justificados por la evidencia de que existen mecanismos causales similares responsables de generar los comportamientos aparentemente similares que se observan. Si el soporte teórico y experimental proporciona esas pruebas, no debería haber ninguna objeción a usar el mismo lenguaje descriptivo para los humanos y no humanos (Mitchell, 2005).

Es generalizado considerar que el Canon de Morgan ha permitido abordar el análisis de la conducta animal de manera sensata para evitar el antropomorfismo superficial que llevó a muchos absurdos en el pasado (Manning y Dawkins, 2012); un antídoto necesario para toda investigación en psicología y comportamiento animal, que ayuda a separar a la “buena ciencia” de lo meramente anecdótico (Fitzpatrick, 2008). El Canon, a su vez, también presenta ciertos problemas: no parece haber consenso en qué quiere decir con “explicación más simple”, y los términos “inferior” y “superior” hoy han sido reemplazados por otros.

¿Qué es una explicación “más simple”? El aprendizaje para Morgan se explica por asociaciones dentro del tipo de conducta de ensayo y error, no por percepción de relaciones como en el caso humano. Distingue entre reacciones innatas y adquiridas en los animales, y admite la imitación como fuente de adquisición de experiencias. Morgan reconoce cierta predisposición en el humano a atribuir estados mentales a animales para predecir sus comportamientos (Morgan, C. L., 1896). Para Allen-Hermanson (2005), la hipótesis más simple es que otros organismos son como nosotros; la hipótesis más simple es solo el antropomorfismo ingenuo. Si me acerco a una fuente para beber agua porque creo que el agua apaga la sed y quiero dejar de tener sed, entonces la inferencia simple es que otros organismos que se acercan a la fuente y beben agua lo hacen por la misma razón.

En cuanto a los términos “superior” e “inferior”, los teóricos contemporáneos generalmente parecen tomarlos para referirse a la relativa sofisticación de los procesos cognitivos que se postulan como explicaciones rivales para el comportamiento animal. Un proceso cognitivo es “superior” en la escala psicológica que otro si es más sofisticado, y “menor” en la escala si es menos sofisticado (Fitzpatrick, 2008). Algunos autores como Shettleworth (2010) sugieren que los psicólogos suelen utilizar “menor” para referirse a las capacidades cognitivas como el aprendizaje asociativo o comportamientos específicos de especies (anteriormente, “conductas innatas”), y “más alto” para referirse a procesos cognitivos distintos al aprendizaje asociativo (como el razonamiento, planificación, o conocimiento).

Uno de los problemas de la propuesta del Canon, en opinión de Heyes (1998), está en que no hay una razón suficiente para suponer que el uso de la teoría de la mente o de procesos mentales en animales, requiere menos inteligencia por parte de estos. Ni la inteligencia ni la simplicidad pueden medirse de una forma razonable.

Aunque el Canon podría tomarse, al menos en su primera formulación, como un principio que imposibilitaría explicar gran parte del

comportamiento animal atribuyéndole al mismo una base mental compleja, en Morgan (1903) agregó una segunda parte a su Canon, donde asegura que no excluye la interpretación de una actividad en particular en términos de procesos superiores, si ya tenemos evidencia independiente de la ocurrencia de estos procesos superiores en el animal en observación:

A esto, sin embargo, hay que añadir que para que el alcance del principio no sea malentendido, el Canon no excluye la interpretación de una actividad en particular en términos de procesos superiores, si ya contamos con pruebas de la existencia de estos procesos superiores en el animal que estamos observando (Morgan, 1903, 59).

Glock (2009) propone reemplazar el Canon de Morgan por el suyo propio que es, según dice, más modesto. Así, mientras Morgan sugería atribuirle capacidades mentales de orden superior a una criatura solamente si esta es la *única* explicación de sus capacidades conductuales, Glock sugiere atribuir capacidades de orden superior a una criatura solo si esta es la *mejor* explicación de tales capacidades. Como aquel, el Canon de Glock reposa en una clasificación gradual o estratificada de las capacidades mentales, que van desde las de orden superior hasta las de orden inferior. No se trata realmente de un canon completamente nuevo; es más bien un debilitamiento de la propuesta de Morgan, que evitaría la invocación por parte de algunos investigadores de logros cognitivos en los animales, apelando a hazañas de conocimiento inverosímiles, para evitar atribuirles facultades mentales superiores.

#### 4. ANIMALES LINGÜALISTAS: SPELKE, DAVIDSON Y DE WAAL

Frente a la pregunta sobre si los animales poseen conceptos, predominan dos posiciones extremas relativas a las mentes animales: afirmar que algunos sí tienen la capacidad de poseer tanto conceptos como creencias, o negar de plano que posean conceptos o creencias. Esta cuestión es la base de una intensa controversia interdisciplinaria: el interés de los filósofos en esta materia se deriva principalmente de la convicción de que poseer conceptos es un factor clave que distingue al humano de los animales no humanos. Esta diferencia cognitiva es explotada para justificar distinciones importantes en el comportamiento moral de los humanos respecto a otros animales. Por su parte, psicólogos (así como los filósofos de la mente) buscan conocer cómo funciona la mente humana y sus diferencias de frente a la de otros animales. Los lingüistas (así como

los filósofos del lenguaje) pretenden investigar si la capacidad de formar conceptos y creencias se limita a los seres humanos y si esta capacidad puede ser vista como la base para el conocimiento de las lenguas. Los investigadores de la conducta animal apuntan a la comprensión de los procesos causales subyacentes en las capacidades cognitivas sorprendentes de ratas, pájaros y monos, y tienen por objeto aclarar cómo estos procesos causales se relacionan con las capacidades cognitivas humanas. Todos estos enfoques presuponen una cierta noción de "concepto" y a menudo conducen a diferentes reclamaciones relativas a una pregunta clave: ¿Los animales no lingüísticos poseen conceptos y creencias? (Newen y Bartels, 2003).

Muchas criaturas no lingüísticas se comportan de maneras que parecen requerir el tratamiento de criaturas pensantes. Pero realmente no tenemos ninguna manera de atribuir pensamientos a las criaturas no lingüísticas que no pase por la cruda analogía de la atribución de pensamientos a las criaturas que utilizan el lenguaje. No tenemos ningún marco teórico para entender el contenido y naturaleza del pensamiento no lingüístico o los mecanismos de razonamiento y la reflexión de la que criaturas no lingüísticas pudieran ser capaces. En ausencia de un marco teórico, las prácticas explicativas dentro de las cuales parece tan necesaria la atribución de pensamientos a las criaturas no lingüísticas permanecen sin una base segura (Bermúdez, 2003).

A causa de la conexión íntima entre el lenguaje y los conceptos humanos, algunos niegan que los conceptos puedan ser atribuidos a otros animales. Donald Davidson quizá sea el autor más citado y reconocido de quienes niegan que los animales sin lenguaje sean capaces de algún tipo de pensamiento.

La convicción de que sin lenguaje plenamente desarrollado es imposible hablar propiamente de mente y no se dispone de algo que pueda llamarse pensamiento, está bastante arraigada en la historia del pensamiento y extendida culturalmente. Si no se posee un lenguaje complejo capaz de someterse a las reglas sintácticas, con orden y estructura, no se puede hablar de mente y/o procesos mentales. Si bien la mente no es un producto exclusivo del lenguaje, solo se puede tener plena seguridad de la existencia de procesos mentales en animales con capacidad lingüística (Diéguez, 2014).

Algunas de las primeras preocupaciones de Davidson como filósofo estaban enfocadas en la relación verdad-proposiciones, alejadas del tema de la cognición animal. Él creía que los animales no son capaces de actitudes proposicionales, por lo que solo pueden tener pensamientos simples,

sacándolos de las principales áreas de interés cognitivo (Davidson, 2001). Su premisa es que toda creencia tiene un contenido proposicional, distante de las capacidades cognitivas de los animales no humanos, al no poseer lenguaje.

Para Davidson, es necesario que un organismo tenga ya el concepto de creencia para que pueda tener creencias, lo que lleva implícito el discernir entre verdad y error de tales creencias, pues el concepto de creencia alberga la posibilidad de que estas sean verdaderas o falsas. Considera que tal organismo debe hacer parte de una comunidad de habla, esto es, poseer y ser capaz de interpretar un lenguaje (Diéguez, 2012).

Nos dice Davidson:

Para que una persona crea que está viendo un gato, debe saber qué es un gato, qué es ver, y sobre todo, debe reconocer la posibilidad, por muy remota que sea, de que se puede estar equivocado. Algunos suponen que los perros pueden tener una creencia aislada, pero creo [...] que los perros no tienen creencias o cualquier otra cualidad proposicional. No hacen juicios. [...] La razón por la que ni los perros o cualquier otra criatura puedan tener creencias aisladas, tales como estar viendo un gato, es que lo que identifica a una creencia, es lo que en términos generales llamamos un contenido proposicional. Así, para tener una creencia sobre un gato, se deben dominar los conceptos que estén implicados en el juicio o creencia (Davidson, 1999a, 8).

Así, no puede haber creencias (ni pensamiento) si no se tiene previamente el concepto mismo de estas, lo que lleva a su vez a la exigencia de la capacidad de usar y entender un lenguaje. El lenguaje es así un requisito indispensable para el pensamiento.

Mi tesis no es entonces –escribe Davidson en ‘Rational Animals’– que la existencia de cada pensamiento dependa de la existencia de una oración que exprese ese pensamiento. Antes bien, mi tesis es que una criatura no puede tener un pensamiento a menos que tenga un lenguaje. Para ser una criatura racional pensante, esa criatura debe ser capaz de expresar muchos pensamientos, pero sobre todo, ser capaz de interpretar el habla y los pensamientos de los demás (Davidson, 1982, 322-323).

Davidson acude al *factor sorpresa*, queriendo mostrar que para tener creencias es necesario tener el concepto de creencia. La sorpresa, escribe, “requiere que yo sea consciente de un contraste entre lo que yo creía y lo que he llegado a creer. Este conocimiento, sin embargo, es una creencia acerca de una creencia: si me sorprende, entonces, entre otras cosas llego a creer que mi creencia original era falsa” (Davidson, 1982, 326). La

sorprende requiere creencias acerca de la corrección de las creencias propias: <Creí que había llegado tarde, pero al ver mi reloj noté que llegué a tiempo>. Sorprenderse implica tener creencias acerca de las creencias.

Como Davidson, Dretske (1995) considera que para tener creencias hace falta poseer el concepto de creencia, por lo que, por ejemplo, un ratoncillo está imposibilitado para tener la creencia de que una persona está tocando el piano. A su entender, no puede haber creencias sin conceptos; toda creencia ha de tener un contenido conceptual. Pero contrario a aquel, cree que puede haber representaciones mentales no conceptuales (como las construidas por experiencias sensoriales, por ejemplo, escuchar el piano): estas representaciones mentales pueden ser extendidas a los demás animales. Así, según Dretske, los animales no humanos sí pueden tener representaciones mentales, aunque no pueden tener creencias.

Davidson explora cómo establecer la evidencia relevante para decidir si un organismo posee o no actitudes proposicionales. Argumenta que la atribución de tales actitudes corresponde justamente a la atribución de racionalidad, considerando que la atribución de creencias es un caso de actitud proposicional. Intenta mostrar que un ser racional es un ser dotado de creencias y, como ya se dijo, para tener creencias se debe contar con el concepto de creencia, lo que depende de manera decisiva de la posesión de un lenguaje. En la medida en que el lenguaje es un proceso social, la racionalidad también lo es.

Como se dijo, Davidson otorga un papel primordial al pensamiento proposicional. Pero cada paso individual del pensamiento proposicional tiende a ser simple en comparación con la complejidad encarnada por las interacciones de pensamiento automático no consciente. El pensamiento proposicional es muy flexible en cuanto a su capacidad para tomar nuevas direcciones, porque es transmisible entre personas. Cualquier nueva dirección se puede construir a través de esfuerzo colectivo. Eso es lo que hace que el pensamiento proposicional sea tan eficaz. Igual que el pensamiento automático, pero actuando con mayor rapidez debido a la transmisión cultural, el *efecto trinquete* enfatizado por Tomasello, dotado de potentes efectos acumulativos en el tiempo.

Si Davidson estuviese en lo cierto, no solo habría que negar la posesión de lenguaje a niños menores de un año, sino que también la posesión de creencias deseos, y diversos estados intencionales (Diéguez, 2012). Por esto Davidson afirma que “hay un gran problema en saber cómo describir los estados mentales de un niño que solo está parcialmente metido en el lenguaje y el tipo de pensamiento que encaja con ellos” (Davidson, 1999b, 305). Para Davidson, tanto en niños en la etapa prelingüística de

desarrollo como en animales no humanos, no hay algo a lo que pueda llamarse pensamiento propiamente.

Elizabeth Spelke (Spelke, 1990; 1994; 1998; Spelke y Tsivkin, 2003) y su grupo de investigadores, tras varios años trabajando con niños recién nacidos y de pocos meses de edad, concluyeron que pueden realizar inferencias de cierta complejidad acerca del comportamiento físico de los objetos. Ellos han logrado identificar algunas expectativas inherentes en bebés de apenas una o dos semanas, midiendo el tiempo que miran una escena en la que esas expectativas fueron o no satisfechas. Cuando las expectativas no se cumplen, los bebés experimentan un sentido de sorpresa y asombro.

Aunque carecen de lenguaje –y por tanto de conceptos y pensamientos según Davidson–, los niños parecen tener algunas ‘creencias’ sobre el mundo. Antes del primer año de edad, saben lo que es un objeto: una unidad física concreta en la que todas las partes se mueven más o menos como una sola, y con cierta independencia de otros objetos. Si tomo la esquina de un libro que está sobre la mesa, los niños esperan que el resto del libro se venga conmigo, pero no la mesa. Si se muestra a un bebé una secuencia del truco en el que una barra que parece ser sólida se mueve hacia atrás y adelante detrás de otro objeto, el bebé abrirá su boca con asombro cuando se elimine dicho objeto y la barra resulte ser dos fragmentos. También saben que los objetos no pueden ir a través de límites sólidos u ocupar la misma posición que otros objetos, y que en general, viajan por el espacio en una trayectoria continua.

Bebés y niños pequeños usan pistas geométricas para orientarse en el espacio tridimensional, navegar a través de habitaciones y localizar “tesoros” ocultos. Al tiempo, los resultados de las investigaciones sugieren que los niños pequeños son bastante malos en el uso de puntos de referencia para encontrar un camino. No es sino hasta los 5-6 años que comienzan a aumentar sus estrategias de búsqueda de pistas como ‘ella ocultó mi juguete en una esquina cuya pared izquierda es de color azul en lugar de rojo’. Estos módulos mentales básicos –representación de objetos y de navegación geométrica–son sistemas al menos parcialmente compartidos con otros animales; por ejemplo, las ratas también navegan por un laberinto a través de la forma, pero no del color.

El control mutuo de la atención y la asignación de recursos en actividades sociales coordinadas no requiere intervención del lenguaje. La coordinación requiere un grado de comunicación, pero perfectamente puede ser no simbólica y, por tanto, no lingüística. Es sabido, por ejemplo, que los bebés humanos se comprometen a partir de una edad muy temprana

en períodos sostenidos de actividad coordinada con sus cuidadores. Este proceso es denominado por Daniel Stern (1991) como "sintonía afectiva". Es un proceso de explorar y comunicar estados emocionales a través de cambios en la expresión facial, vocalizaciones y gestos.

Para explicar lo que sucede en niños y animales, Davidson (1999a) recurre a la "triangulación": ambos aprenden a correlacionar cambios y reacciones de otros con cambios u objetos del mundo ante los cuales ellos también reaccionan, como lo hacen los peces de un banco cuando reaccionan al movimiento de otros de su grupo; esto es, interacción a tres bandas entre dos individuos y el mundo. La triangulación es, a su entender, el mecanismo mediante el cual llegamos a tener conocimiento de la realidad. Además, nos proporciona también la única explicación de cómo la experiencia da un contenido específico a nuestros pensamientos. Sin las demás personas con las cuales compartir respuestas a un entorno mutuo, no existe una respuesta a la pregunta de qué es a lo que estamos respondiendo en el mundo.

Para Davidson (2003), el conocimiento intersubjetivo es el conocimiento de las mentes de otros. Pero como no podemos conocer la mente de otras personas directamente, es necesario acudir al lenguaje y a las acciones del agente para poder estar al tanto de sus actitudes proposicionales. Cuando observamos acciones de otro o prestamos atención a sus palabras, es posible llegar a aprehender sus creencias, deseos e intenciones, porque "todo el que entiende el lenguaje puede reconocer aseveraciones y sabe que quien hace una aseveración se representa a sí mismo como alguien que cree lo que dice" (Davidson, 2003, 285). Es igual con las acciones: cuando una persona actúa de cierta manera, asumimos que lo hace obedeciendo a ciertas actitudes proposicionales. Consideramos las acciones de las personas como intencionales y racionales. Así, para Davidson se puede decir que a partir de las acciones y manifestaciones lingüísticas de otras personas conocemos sus actitudes proposicionales, y conocer sus actitudes proposicionales es conocer lo que hay en la mente de una persona. La intencionalidad y 'teoría de la mente' están, de esta manera, también supeditadas al lenguaje.

Empero, algo similar podría suceder en el caso de otros animales. James Gould (Gould, 1982; Dyer y Gould, 1983) sugiere que las abejas pueden tener un mapa cognitivo de la información que han aprendido, mapa que utilizan para comunicarse. En varios de sus experimentos, trasladó un suministro de jarabe de azúcar un 25% más lejos de una colmena diariamente. Las abejas se comunican entre sí la ubicación. Después se coloca el jarabe de azúcar en un barco anclado en el centro de un pequeño lago. Cuando las exploradoras regresaron a la colmena para comunicar



su hallazgo, otras abejas se negaron a ir con ellas en busca de esta fuente de alimento, no esperando encontrar comida en el medio de un lago, a pesar que con frecuencia volaban sobre el mismo para llegar a fuentes de polen en la orilla opuesta. En otra prueba, Gould atrajo algunas abejas a un plato de néctar artificial, luego poco a poco lo trasladó más lejos de la colmena después de que se acostumbraron a esta fuente de alimento. El equipo marcó a las abejas entrenadas, los colocó en un frasco oscuro, y los trasladó a un lugar donde la colmena todavía era visible, pero no el plato. Cuando se soltaron una a una, las abejas parecían desorientadas por unos segundos, y luego volaron directamente hacia el plato. 73 de 75 abejas llegaron en unos 28 segundos. Al parecer, lograron esta hazaña mediante la elaboración de una nueva ruta de vuelo sobre la base de un mapa cognitivo de los lugares visibles.

Lo arriesgado de la explicación de Gould (cfr. Gould, 2002) está en que sugiere que las abejas manejan los conceptos de simetría, 'igual' y 'diferente'. El problema aquí es que una cosa es mostrar que los animales puedan discriminar entre objetos –es decir, que son capaces de categorizar–, y otra es sugerir que poseen y utilizan conceptos abstractos.

Que los animales puedan discriminar no puede tomarse como razón suficiente para atribuirles conceptos pues, teniendo en cuenta que prácticamente todos los seres vivos pueden hacerlo, por ejemplo, entre depredador/presa o nutritivo/tóxico, tendríamos que atribuir generosamente a todos los organismos la posesión de conceptos, con lo que el concepto mismo de 'concepto' quedaría vacío de contenido (Diéguez, 2014).

Los estudios encabezados por Paul Quinn (Quinn *et al*, 2003; Quinn y Tanaka, 2007), indican que los bebés prelingüísticos cuentan con capacidad de categorización con cierto nivel de abstracción, mostrando que aun en niños muy pequeños existe la capacidad de formar una representación categórica de una relación espacial. Sin embargo, los bebés no proporcionan evidencia hasta los 9 meses de edad de poder generalizar la relación espacial entre nuevas formas; 6 meses parece ser la edad más joven en la que demuestran capacidad de formar una representación categórica abstracta a través de formas desconocidas, y hasta ahora han demostrado esta capacidad solo respecto a las relaciones espaciales de arriba/debajo.

Aunque está claro que estas categorizaciones no son propiamente conceptos, parece que los conceptos que manejamos los adultos son un enriquecimiento en el contenido de esas categorizaciones básicas, donde el lenguaje, aunque no indispensable, sí es de gran ayuda. La evidencia confirma, según Quinn y su equipo, que las representaciones para diferentes relaciones espaciales surgen en diferentes puntos durante

el desarrollo, y sugiere que cada representación se somete a su propio período de desarrollo de lo concreto a lo abstracto.

Allen (1999) considera que la estrecha relación entre lenguaje humano y conceptos humanos hace que sea altamente cuestionable la atribución de conceptos a animales no humanos. Afirma que es importante establecer cuándo estaríamos dispuestos a considerar que cierto organismo posee un concepto, sin presuponer que solo sea cuando tiene la capacidad de expresarlo lingüísticamente. Propone así un enfoque de tripartito para la atribución de conceptos a animales. Dicho enfoque va más allá de las pruebas habituales de discriminación mediante búsqueda de pruebas para la auto-monitorización de los errores de discriminación. Tal evidencia puede recogerse sin depender de la lengua. Argumenta que la capacidad para la detección de errores solo puede explicarse mediante atribución de una especie de representación interna identificable razonablemente como concepto.

Según su enfoque, es razonable atribuir a un organismo dado un concepto de X si, primero, ese organismo discrimina sistemáticamente algunos Xs de algunos que no lo son; segundo, si el organismo es capaz de detectar algunos de sus errores de discriminación entre Xs y no-Xs; y tercero, si el organismo en cuestión logra aprender a discriminar Xs de no-Xs como consecuencia de su capacidad anterior (Allen C., 1999). En resumen: podemos considerar que un animal posee un concepto si es capaz de discriminar Xs de no-Xs y aprender a mejorar la práctica clasificatoria por ensayo y error, esto es, aprendiendo de sus propios errores (Diéguez, 2012).

Según este enfoque, aunque todos los animales cumplen el primer requisito de Allen –pues, como se dijo, todos pueden discriminar–, la capacidad de conceptualización solo sería posible en animales cuya inteligencia les permita cumplir con los tres requisitos. La distancia entre la mera discriminación y la conceptualización está entonces en la facultad de detectar y reconocer los errores en la clasificación y aprender de ellos. Lo importante es que el organismo detecte y corrija sus errores, aunque no posea la noción de error ni prevea la posibilidad de este, lo que –contra Davidson–, no exige posesión de lenguaje.

Aunque la propuesta de Allen pretende ser restrictiva, no deja de ser lo bastante laxa como para que sean atribuidos fácilmente conceptos a animales no lingüísticos. Un ejemplo puede ilustrarlo: pocas horas después del nacimiento, ya los pollitos están dotados de movimientos reflejos convenientemente apropiados. Corren de un lado a otro, picotean los objetos que se les ofrece y los alcanzan con seguridad. Al principio, los pollitos picotean cualquier cosa, ya sea los caracteres impresos en una hoja de

papel, granos de arena, sus propias patas y excrementos. En este último caso, de inmediato rechazan el objeto de mal sabor, sacuden la cabeza y limpian su pico frotando contra el suelo. Lo mismo cuando pican una abeja o una oruga de sabor desagradable. Pronto cesan de picotear los objetos que no les son agradables.

En este sencillo ejemplo se cumplen los tres requisitos propuestos por Allen, con lo que se evidencia que su filtro no impide atribuir fácilmente el concepto de *alimento* a un pollo.

Pero algunos autores como Thompson y Oden (2000), consideran el enfoque en tres partes de Allen demasiado exigente. Afirman que un concepto es solo el conocimiento necesario para realizar una categorización. Basta con solo ser capaz de realizar una clasificación basándose en la comprensión de ciertas relaciones entre objetos, de manera que se puedan trazar analogías entre ellos: si se clasifica juntos un plátano y una manzana y se excluye de la misma a un peluche porque los dos primeros son comestibles y el otro no, entonces se tiene de alguna manera el concepto de alimento, aunque no se exprese lingüísticamente (Diéguez, 2012).

De acuerdo con esta propuesta, los conceptos son principios de discriminación, con lo que poseer conceptos es poseer la capacidad de reconocer o discriminar entre diferentes tipos de objetos. Esta posición es bastante generosa, pues no se requiere de mucho esfuerzo para atribuir conceptos a los animales en cautiverio o en estado salvaje. Aquí parece propicio otorgar la razón a Davidson: conceptualizar requiere algo más que discriminar.

Herrnstein (1990) estudia algunos experimentos sobre la categorización de los estímulos visuales en animales no humanos. Los resultados sugieren, según Herrnstein, una clasificación de las competencias categóricas en cinco niveles ascendentes: (i) discriminación simple; (ii) categorización por memorización; (iii) categorización abierta basada en similitud perceptiva; (iv) conceptualización; y (v) uso de relaciones abstractas. Considera que a los animales no humanos les resulta fácil categorizar hasta el cuarto nivel, que es el nivel de los conceptos. Con dificultad, a veces podrían ser inducidos a elevarse hasta el nivel de las relaciones abstractas. Es en el nivel de las relaciones abstractas donde existe una gran brecha entre las categorizaciones humanas y la categorización por otros animales.

Vauclair (2002) ha estudiado en detalle las categorizaciones hechas por primates no humanos, prestando especial atención no a la mera discriminación, sino a la posible capacidad de formar clases asociativas en las categorizaciones, lo que implica cierto nivel de abstracción, en

comparación a la categorización puramente perceptiva. Los babuinos, por ejemplo, clasifican entre cosas comestibles y no comestibles, incluso mediante imágenes. Más aún, no solo clasifican una manzana como similar a otra manzana por su forma o color, sino que la clasifican como similar a un plátano al ser ambos alimentos. La discusión en estos casos está, sin embargo, en si es posible que tales animales lo logren basándose en ciertas percepciones que no puede percibir el experimentador (Diéguez, 2014).

Pero hay resultados todavía más curiosos. Seis orangutanes (*Pongo abelii*) de diferentes edades debían clasificar fotografías de orangutanes por medio de una pantalla táctil frente a otras con imágenes de diversos tipos de primates, fotografías de primates frente a fotografías de no-primates, y fotografías de animales frente a otras de no-animales. Cada nivel requería un mayor nivel de abstracción. Al final, los orangutanes tuvieron más dificultad en el nivel 3 (imágenes de animales frente a no-animales). En el nivel dos (imágenes de primates frente a imágenes de no-primates), sus resultados fueron buenos. Los resultados sugieren que los orangutanes pueden aprender conceptos en cada nivel de abstracción; la mayoría de estos sujetos aprendieron rápidamente la discriminación de nivel intermedio (Vonk y MacDonald, 2004). Sin embargo, los investigadores reconocen no obtener conclusiones definitivas, pues es imposible determinar si el aprendizaje de la discriminación anterior interfirió o facilitó el aprendizaje de la discriminación más abstracta.

Dos años antes de esta prueba, ya se había realizado una similar con Zuri, una joven gorila del zoológico de Toronto. En el experimento también diseñado por Vonk y MacDonald (Vonc y MacDonald, 2002), Zuri fue capaz de clasificar fotografías de seres humanos frente a otras con gorilas y orangutanes, incluyendo acertadamente la fotografía de un gorila albino entre las de los gorilas. Esta clasificación, sin embargo, podría estar basada en rasgos perceptivos. Pero Zuri también fue capaz de clasificar correctamente con criterios más abstractos fotografías de animales frente a fotografías de no-animales, y fotografías con imágenes de animales frente a imágenes de alimentos. Los resultados fueron más pobres al clasificar de acuerdo al nivel intermedio de abstracción. Vonk y MacDonald sugieren, aunque tampoco de manera concluyente, que en los gorilas la categorización tiene una base conceptual y no solo perceptiva.

Los resultados no se limitan solo a mamíferos. Algunos estudios con aves como palomas muestran que pueden manejar conceptos abstractos como "esfericidad", efectuar inferencias sobre transitividad y resolver problemas basándose en reglas abstractas (Diéguez, 2005); (Delius y Godoy, 2001).

Tal parece que lo importante es el grado de abstracción en la capacidad discriminatoria, esto es, detectar si existen animales que puedan clasificar objetos no simplemente basándose en las propiedades perceptivas, sino que también en las propiedades funcionales o relacionales de los objetos.

En este extremo del debate sobre los procesos mentales animales se encuentran los que afirman que los pensamientos de los animales difieren de los pensamientos humanos solo en grado, debido a sus diferentes entradas perceptivas. Las ostras, por ejemplo, no tienen pensamientos sobre bicicletas, simplemente porque no pueden percibir las bicicletas. Glock (2015) resalta una posición intermedia entre los extremos *diferencialistas* (que niegan todo tipo de pensamiento animal) y *asimilacionistas* (que ponen el pensamiento de los animales a la par con el pensamiento humano). Una coalición algo rara entre el sentido común y Wittgenstein: los animales son capaces de tener pensamientos de un tipo sencillo, es decir, aquellos que se pueden expresar en la conducta no lingüística.

Oponiéndose a la posición de Davidson, Glock se decanta por una variedad de esta posición intermedia: "Podemos adscribir pensamientos y conceptos a los animales, pero estos quedan restringidos a un tipo simple, ya que solo los pensamientos simples pueden ser manifestados identificablemente en un comportamiento no lingüístico" (Glock, 2009, 83). Considera que hay un género de pensamiento perceptivo que no requiere conceptos, y que la posesión de estos no está ligada al lenguaje, sino a la capacidad de hacer discriminaciones sujetas a evaluación normativa. "El comportamiento de las criaturas no lingüísticas no es siempre explicable únicamente haciendo referencia a los imperativos biológicos inmediatos" (p. 122), por lo que según Glock, Davidson se equivoca al sostener que las criaturas no lingüísticas solo tienen disposiciones y carecen de capacidades. Tanto los bebés humanos como los primates superiores son capaces de actuar voluntariamente, pues pueden abstenerse de llevar a cabo una acción particular, ya sea buscando conseguir su objetivo de otro modo o renunciando a él.

De este modo, la posición intermedia asumida por Glock no requiere que la posesión de conceptos dependa necesariamente de la posesión de lenguaje, sino más bien de un comportamiento discriminatorio lo suficientemente complejo y flexible como para estar sujeto a valoración normativa.

Tal posición considera que no existe razón que nos obligue a afirmar que los animales no pueden poseer conceptos. Incluso en caso de haberla, no es suficiente como para negarles la posibilidad de tener pensamientos, dada la posibilidad de creencias de tipo perceptivo. El que los

animales tengan o no conceptos de tipo simple, según Glock, no depende de la posesión de lenguaje, sino de hasta qué punto sus discriminaciones están gobernadas por reglas y así, sean intencionales.

Otro asunto es la capacidad de razonar y comprender la causalidad detrás de los fenómenos más simples. Algunos investigadores (Call, 2006; Tomasello, Call, y Hare, 2003), han obtenido algunas conclusiones: simios (y posiblemente otros animales) son en realidad bastante buenos en la comprensión y el razonamiento acerca de ciertas propiedades físicas del mundo, mientras que, al mismo tiempo, son bastante malos al asociar estímulos arbitrarios y respuestas. En otras palabras, si dos estímulos tienen una conexión causal (como cuando el alimento suena dentro de una taza al sacudirlo), los simios obtienen mejores resultados que si los estímulos tienen una relación arbitraria (como cuando un ruido sin relación indica alimento). Además de la capacidad de decidir sobre fenómenos físicos, Call argumenta que los monos (y otros animales) también tienen algún acceso a la comprensión de problemas. Tienen capacidad metacognitiva que les permite saber que saben o no saben algo: los monos *rhesus*, como los delfines, saben cuándo no están seguros de algo o cuando lo han olvidado. Los chimpancés saben cuándo no han visto algo determinado con anterioridad, cuando sus compañeros han visto o no algo y saben además cuando están seguros o no de algo. De este modo, cree Call que el razonamiento y la reflexión no pueden ser criterio de la singularidad humana, como pretendía Descartes. Más bien, estas habilidades pudieron haberse desarrollado (o coevolucionado) en otros animales también, porque ello les permite solucionar problemas en el mundo de una manera más eficiente.

Algunos señalan que, si bien los animales pueden tener representaciones mentales, estas no tienen el carácter de representaciones abstractas desvinculadas de la situación que las suscita. Son este tipo de representaciones las que permiten a los humanos la planificación a largo plazo, facultad de la que carecerían los demás animales. Sin embargo, existe evidencia de que bonobos y orangutanes transportan los elementos adecuados que posteriormente utilizarán para alcanzar la comida (Mulcahy y Call, 2006). Esta capacidad para conservar y transportar elementos que luego emplearán, puede interpretarse como una capacidad para planear con antelación el desarrollo de una tarea futura. En el mismo sentido, un equipo de investigadores dirigidos por Elisabetta Visalberghi (Visalberghi y Fragaszy, 2013), documentó cómo un grupo de monos capuchinos utilizan piedras para cascar nueces previamente seleccionadas por estos 'sabiendo' cuál es la mejor para cada trabajo. Toman varias piedras y las

van probando hasta decidirse por la más pesada, la cual es la más resistente para abrir las nueces de palma. Más diciente al respecto es la evidencia obtenida por Pruetz y Bertolani (2007): un grupo de chimpancés de Fongoli prepara lanzas que posteriormente serán utilizadas para la caza de primates más pequeños.

Además de la capacidad de detectar errores y corregirlos (dos de los filtros de Allen), otro tema importante al hablar de procesos mentales en animales no humanos es también el evitarlos. En un experimento con siete gorilas, ocho chimpancés, cuatro bonobos y siete orangutanes, Call (2010) examinó que estos no solo detectan y corrigen sus errores, sino que, además, detectan cuando están equivocados. En pruebas donde se incrementa el tiempo de espera para dar una respuesta que consiste en seleccionar un tubo que oculta comida o se aumenta el valor de la recompensa, por ejemplo, con comida más apetitosa, se aseguran con más cuidado de no fallar en la respuesta.

Al descartar que recordaban mejor la ubicación de la comida menos apetecible, los sujetos buscaban tener un mayor grado de certeza en la elección cuando en caso de equivocarse la pérdida era de mucha importancia para ellos. El experimento básicamente consistió en colocar dos tubos, uno con comida y otro sin ella (o con comida no muy apetitosa). A los simios se les permitía hacer solo una elección; si acertaban se quedaban con la comida. Los resultados arrojaron que, cuando la comida era muy apetitosa para ellos, por ejemplo, una banana, los simios miraban el interior del tubo varias veces para constatar su presencia, pese a que poco antes habían visto al experimentador meterla, acción que no hacían cuando la comida no era muy agradable.

Este estudio sugiere que la toma de decisiones en grandes simios no es muy diferente de la nuestra cuando poco tiempo antes de viajar, por ejemplo, constatamos varias veces tener los documentos en regla para tal efecto, pese a 'estar seguros' de llevarlos completos (Call, 2010).

Los nuevos resultados –escribe Call–, [...] sugieren que los animales no humanos pueden poseer algunas habilidades metacognitivas [...] Por lo menos, los hallazgos actuales deberían servir para desafiar a los defensores de la posición contraria a producir un nuevo conjunto de explicaciones no metacognitivas que den cuenta de los resultados actuales (Call, 2010, 699-670).

Este dilema sobre si los demás animales, principalmente los primates, son capaces de procesos mentales, podría resumirse en optar por una de dos posiciones opuestas: la *economía cognitiva* sustentada en el Canon de Morgan y la *economía evolutiva*. La primera, como ya se dijo,

sugiere no invocar procesos mentales superiores si un fenómeno puede explicarse a través de procesos inferiores (el condicionamiento, por ejemplo, podría reemplazar la intencionalidad). La economía evolutiva, por su parte, otorga importancia a la filogenia compartida. Considera que, si dos especies con un vínculo de parentesco cercano se comportan de la misma forma, es muy probable que los procesos mentales subyacentes sean los mismos. Pero, si no tenemos mayor problema en asumir que un posible comportamiento similar de un chacal, un coyote y un lobo puede ser ocasionado por las mismas causas, ¿por qué buscar causas diferentes cuando ello ocurre entre bonobos, chimpancés y humanos? (De Waal, 2007).

El detalle podría estar en que, como se dijo anteriormente, una genealogía común no significa que no haya diferencias cualitativas entre los rasgos exhibidos por especies relacionadas. El lenguaje simbólico marca una diferencia considerable entre nosotros y otros animales, incluso los más próximos en parentesco a nosotros.

La posesión de un lenguaje sofisticado con una sintaxis compleja, además de incrementar nuestras posibilidades de comunicación, nos permitió, como afirma Daniel Dennett (2006), ser inquisitivos y no solo curiosos y contemplativos; nos permitió preguntarnos acerca del mundo y su funcionamiento. “No se puede olvidar que la aparición del lenguaje ha hecho de la mente humana algo sumamente peculiar que excede en sus capacidades a cualquier cosa conocida en los otros animales. Nuestro lenguaje modifica y potencia nuestras capacidades cognitivas de una forma extraordinaria” (Diéguez, 2012, 323-324).

Considerando la relación de proximidad evolutiva entre humanos y chimpancés, no es absurdo pensar que los paralelismos externos del comportamiento tengan sus equivalentes internos; es decir, que existen algunos aspectos compartidos entre especies en los mecanismos biológicos que gobiernan el comportamiento y en la experiencia subjetiva correspondiente. Expresiones faciales, movimientos y la postura adoptada que acompaña ciertos comportamientos en chimpancés, por ejemplo, refuerzan esta conjetura (Wright, 2007). Pero nuestras conclusiones sobre tales paralelismos deben estar guiadas y justificadas por evidencia de que existan mecanismos causales similares responsables de generar los comportamientos aparentemente similares que se observan. Si la evidencia analizada críticamente proporciona las pruebas, no debería haber mayor problema en usar el mismo lenguaje descriptivo para dar cuenta de ciertos comportamientos en humanos y otros animales, sobre todo los más cercanos en parentesco.



#### 4.1. *La teoría de la mente*

Las investigaciones sobre las habilidades mentales de los animales no humanos rebosan de controversia. Tales habilidades pueden ser interpretadas al menos desde dos puntos de vista: (i) la hipótesis ‘conductista’, que afirma que los animales no humanos aprenden acerca de las regularidades del mundo o de las conductas observables de los demás a través de mecanismos de bajo nivel similares al condicionamiento pavloviano, sin ninguna capacidad de razonar acerca de la relación causal entre esos comportamientos; y (ii) la hipótesis ‘mentalista’, que propone que los animales no humanos (al menos algunos) tienen estados mentales, tienen ciertas capacidades de razonamiento, así como ciertos estados intencionales. Los hay que incluso consideran que algunos animales atribuyen estados mentales a otros y razonan sobre el papel causal desempeñado por los estados mentales de una manera más o menos análoga a la forma en la que lo hacen los humanos (Penn y Povinelli, 2013).

En los debates sobre cognición animal, el problema de atribuir conceptos, creencias y deseos a los grandes simios es bastante habitual. Pero igual que en el debate específico sobre la atribución o no de pensamientos y conceptos a los animales no lingüísticos, las afirmaciones sobre concederles o negarles intencionalidad a las conductas de los animales no humanos divide las opiniones de los investigadores.

José Luis Bermúdez (2003) por ejemplo, considera que los pensamientos de criaturas no lingüísticas pueden tener estructura composicional que tienen componentes distinguibles, que pueden figurar en otros pensamientos. Poseen contenidos determinados, y es a menudo posible, contra afirmaciones de muchos filósofos, identificar y definir con exactitud razonable, la forma precisa en la que criaturas no lingüísticas están pensando en su entorno. Sin embargo, cree que hay límites a la gama de pensamientos que las criaturas no lingüísticas pueden tener, porque ciertos tipos de pensamiento (todos los que implican ascenso intencional o pensar los pensamientos) requieren un vehículo lingüístico.

Empero, en el proceso de “sintonía afectiva” propuesto por Stern y explicado más arriba, los estados emocionales se comunican cuando se comparten. La participación de los infantes en este tipo de comunicación algunas veces es intencional (como, por supuesto, la del cuidador), y el proceso de sintonía afectiva es un tipo de actividad coordinada en el que hay control mutuo de asignación de recursos. No obstante, el proceso no es de comunicación simbólica, porque las expresiones faciales y vocalizaciones no son símbolos de los estados emocionales que están siendo comunicados.

Michael Tomasello es tal vez uno de los investigadores más citados para sostener que los animales no-humanos tienen intencionalidad, aunque no con el mismo grado de refinamiento que los humanos. Escribe que a pesar de algunas observaciones que sugieren que algunos primates no humanos en algunas situaciones son capaces de entender a sus congéneres como agentes intencionales y de aprender de ellos en formas que se asemejan a algunas formas de aprendizaje cultural humano, el peso abrumador de la evidencia empírica sugiere que solo los seres humanos entienden a sus congéneres como agentes intencionales como ellos mismos, por lo que solo los seres humanos participan en el aprendizaje cultural (Tomasello, 1999).

Es importante una distinción clara entre el concepto de mente y el de teoría de la mente. Como vimos en el apartado anterior, algunos autores (Glock, 2009) defienden que los animales no humanos, al menos algunos, pudieran tener mente, en tanto son capaces de tener intenciones, deseos y creencias, aunque no de carácter proposicional, mientras otros, como Davidson, lo niegan. Pero tener mente no implica necesariamente poseer una teoría de la mente, esto es, tener la capacidad de teorizar sobre la mente de otros, sobre sus deseos, objetivos o creencias.

Se puede identificar la teoría de la mente calibrando su intensidad y llegar así, por ejemplo, a un nivel dos, en el que no solo nos representamos lo que puede estar pensando otro individuo, sino lo que ese individuo piensa que estamos pensando nosotros. Podemos ir a un nivel tres y adivinar lo que el otro piensa que estamos pensando nosotros sobre lo que él está pensando y cómo esta consideración influye en su pensamiento, de manera que, cuanto más se sube de nivel, más potente será la teoría de la mente subyacente (Castrodeza, 2009).

Hasta hace algunos años, no se contaba con una prueba empírica que sugiriera que los demás animales, además del humano, poseyeran una teoría de la mente, pero ciertos experimentos parecen indicar una respuesta positiva. Recordando el Canon de Morgan, cualquier intento de los que se habían hecho para sugerir que los demás animales efectivamente podrían atribuir estados mentales a otros, podía ser explicado por mecanismos más simples.

En este apartado se presentarán diversos ejemplos de experimentos realizados por algunos de los investigadores en el campo de la etología cognitiva. Tales ejemplos no pueden ser considerados como “meramente anecdóticos”, en tanto los mismos son el resultado de experimentos controlados. Sin embargo, como se ha dicho en líneas anteriores, aun cuando algunos ejemplos pudieran resultar de la evidencia anecdótica, no carecen de valor si son analizados de manera crítica. Por otro lado, aunque

haya muchos ejemplos y experimentos que sugieren que efectivamente hay pensamiento en animales no humanos, se debe ir con cuidado a la hora de tomar partido en el debate: nada nos garantiza que no se puedan hacer otras lecturas de los resultados obtenidos en tales experimentos.

Además, nuestra labor en esta área desde la filosofía –como bien afirma Glock– “no consiste en recabar nuevos datos empíricos sobre el comportamiento animal, sus causas neurológicas o sus orígenes evolutivos, sino aclarar en qué consiste la posesión de diversas propiedades mentales y, por tanto, en clarificar bajo qué condiciones se pueden atribuir tales propiedades a los organismos” (Glock, 2009, 61-62). En efecto, la investigación empírica sobre lo que pueden o no pueden hacer otros animales resulta beneficiosa solo en la medida en que tengamos claridad sobre qué condiciones deben ser satisfechas para concluir que ciertos animales poseen o no algo que pueda considerarse mente, pensamientos, creencias o intenciones.

Basado en los diversos experimentos que ha realizado, Call (2001) afirma que los chimpancés cuando siguen la mirada de otros chimpancés o de humanos para ver el objeto al cual estos prestan atención, entienden que sus informantes miran a algo específico en una localización particular. Esto puede interpretarse como una señal de que comprenden que la atención de los otros se dirige a cierto lugar de manera intencional. De esta manera podría considerarse que los chimpancés pueden entender que otros están interesados en objetos o hechos que no están al alcance de su propia vista, pero sí a la de los otros, lo que podría interpretarse como una cierta comprensión del otro como agente intencional.

Pero no todos deducen conclusiones tan generosas. Povinelli (Povinelli, Bering, y Giambrone, 2000), quien ha realizado experimentos en el mismo sentido, concluye que los chimpancés no son capaces de entender el sentido referencial de una mirada. Él y sus colaboradores, consideran que lo único que se puede establecer de tales experimentos es que el seguimiento de la mirada es una respuesta automática resultado de determinados factores que el experimentador no puede asociar con que los chimpancés se representen hacia dónde dirigen su atención los otros. Los chimpancés no comprenden que el otro tiene una experiencia mental visual diferente a la suya, solo han aprendido que si miran en la dirección que el otro lo hace, podrían encontrar algo de su interés.

Afirma Povinelli que el campo de la psicología comparada se basa en dos supuestos. En primer lugar, se supuso que la introspección podría proporcionar el conocimiento fiable sobre la conexión causal entre los estados mentales específicos y comportamientos específicos. En segundo lugar, se supuso que en aquellos casos en los que otras especies exhiben

comportamientos similares a los nuestros, las causas psicológicas eran similares. Considera que este argumento por analogía es deficiente con respecto a los estados mentales de segundo orden. Al centrarse en la cuestión de cómo conciben otras especies la atención visual, y, en particular, si en los chimpancés la observación de la mirada del otro significa que posee estados mentales internos, llega a la conclusión de que los chimpancés no razonan de esta manera, y de hecho, no existe una gran razón para suponer que albergan representaciones de los estados mentales en general.

Pero cinco años después del experimento de Povinelli, los psicólogos Jonathan Flombaum y Laurie Santos (Flombaum y Santos, 2005) consideraron, basándose en sus propios experimentos con monos rhesus (*Macaca mulatta*), que estos poseen la capacidad para deducir lo que otros están percibiendo, basándose en la dirección que están mirando, lo que significa que razonan sobre la percepción visual de otros. Los experimentos consistieron en que los macacos debían robar una fruta a dos sujetos humanos que estaban frente a ellos, uno de los cuales miraba la fruta y otro no. El resultado fue que los macacos, en su mayoría, robaban la fruta que no estaba siendo vigilada por el sujeto humano. Concluyen así Flombaum y Santos (2005, 447) que:

Los macacos rhesus son capaces de utilizar la dirección de la mirada de otro individuo para determinar qué ese individuo puede y no puede ver. Nuestro trabajo se basa en la idea de que los primates muy probablemente pueden exhibir capacidades sofisticadas de 'teoría de la mente' en los escenarios experimentales que imitan las situaciones naturales para las que estas capacidades han evolucionado.

Pero pese a la gran cantidad de experimentos que se habían realizado, incluyendo el citado con monos rhesus realizado por Flombaum y Santos, no se habían obtenido resultados que pudieran ser considerados la 'prueba definitiva' de que otros animales, además de los humanos, tienen una Teoría de la Mente, esto es, detectar sin ambigüedades que un animal, cualquiera que sea, atribuya creencias falsas a otro individuo. No se había detectado que un animal predijera qué hará otro individuo bajo el supuesto de que este tiene una creencia falsa. Se había concluido que un chimpancé puede entender que otro desconoce algo, pero no que crea algo que es falso. Mientras que los niños de 6 años de edad comprenden los dos estados mentales, los chimpancés solo parecían entender el primero, conocimiento-ignorancia, pero no la falsa creencia (Kaminski, Call, y Tomasello, 2008). Comprender que el comportamiento de un congénere depende de la información que posee o, más aún, de sus creencias, que

pueden ser falsas o, simplemente, distintas de las que uno mismo alberga, constituye un nivel psicológico más sofisticado (Colmenares, 2005).

Sin embargo, un reciente experimento con chimpancés, orangutanes y bonobos, sugiere que estos sí anticipan que otros individuos van a actuar de acuerdo con creencias falsas. Christopher Krupenye y su equipo (Krupenye *et al*, 2016) pusieron a 14 chimpancés, 9 bonobos y 7 orangutanes a observar una dramatización y monitorizaron el movimiento de sus ojos para evidenciar que estaban siguiendo la escena en la que una persona disfrazada de simio le robaba una piedra a una persona y la escondía en una caja, después esa persona salía de escena y quien estaba disfrazado cambiaba la piedra de caja y finalmente se la llevaba. Cuando regresaba el individuo a la escena a buscar la piedra, los primates (a sabiendas que la piedra no estaba en ninguna de las cajas) fijaban su mirada en la primera caja conscientes de que el hombre creía que estaba allí escondida. De este modo, concluyen los investigadores, los simios se anticiparon a la acción del hombre, motivada por una falsa creencia.

Un problema importante de las pruebas sobre la teoría de la mente que se practican en simios es la interpretación de los resultados negativos. Algunas veces, los sujetos se encuentran en situaciones bastante inusuales, por ejemplo, con personas con los ojos vendados. Igual que los humanos, los simios son muy sensibles al lenguaje corporal, por lo que un experimentador desprevenido podría confundirlos o alterarlos. Además, las reglas de contacto visual de los simios y humanos son diferentes: los simios, en lugar de mirar fijamente a alguien, optan por vigilar a sus compañeros mediante la visión periférica y con miradas rápidas, poco perceptibles (De Waal, 1997).

Otro de los indicios para estudiar la mayor o menor cognición en los animales en relación con la intencionalidad, es su hipotética capacidad de engañar. Algunos experimentos han permitido sugerir que los chimpancés pueden intentar engañar intencionalmente a un ser humano.

Al respecto escribe Frans de Waal:

El verdadero engaño, puede definirse como la proyección deliberada, para el aprovechamiento de uno, de una imagen falsa de una conducta, de una información o de una intención anterior. En su sentido más absoluto, requiere la conciencia del modo en que se transmiten las acciones que uno realiza y de cómo el mundo exterior puede interpretarlas (De Waal, 1997, 102).

Algunos chimpancés no solo saben qué puede ver o no un ser humano u otro chimpancé, sino que utilizan esa información para conseguir sus fines. En algunas ocasiones se ha documentado cómo se esconden, a

través de una ruta oculta a la vista de los humanos, para conseguir un poco más de comida (Hare, Call, y Tomasello, 2006) y, en un contexto competitivo, los chimpancés saben que si uno de los suyos ha visto donde está la comida, entonces sabe dónde se encuentra la comida (Hare, Call, y Tomasello, 2001).

Mathias Osvath y Elin Karvonen (Osvath y Karvonen, 2012), documentaron cómo un chimpancé en cautiverio llamado Santino, juntaba piedras que extraía de una isla artificial de su recinto en el zoológico de Furuviik, en Suecia, para luego lanzarlas a los visitantes. Los científicos han detectado que Santino es más previsor e innovador de lo que se pensaba. Encubre sus 'depósitos de armas' detrás de troncos y rocas, lo que muestra que los chimpancés son capaces de planear con más complejidad de lo que se conocía. En un momento dado, después de un intento de arrojar las piedras, un guía turístico dejó al chimpancé solo durante horas sin visitantes. Cuando el guía y un grupo de visitantes regresaron, Santino actuó indiferente mientras mantenía los proyectiles y caminaba hacia el grupo; su aspecto no sugirió intenciones de lanzar. El chimpancé se detuvo y cogió una manzana que flotaba en el agua, de la que tomó un bocado mientras continuaba acercándose a los visitantes, y solo dentro del alcance, hizo un tiro repentino y desprevenido al grupo.

En un estudio comparativo sobre la conducta de los primates, Richard Byrne y Nadia Corp (Byrne y Corp, 2004), descubrieron una relación directa entre el tamaño del cerebro y el carácter furtivo. Vieron que cuanto mayor era el volumen promedio de la neocorteza de los primates, mayor era la posibilidad de que el mono o simio protagonizara una hazaña como, por ejemplo, que cuando un joven babuino era perseguido por su madre furiosa, decidida a castigarlo, el animal interrumpía repentinamente su marcha, se erguía y comenzaba a mirar el horizonte con atención, hecho que distraía a todo el grupo y los incitaba a prepararse para la llegada de intrusos inexistentes; o que en un grupo de babuinos en las selvas del Sudán, un pequeño, habiendo descubierto que una hembra había encontrado una rica fuente de alimento, para alejarla y aprovechar él mismo el hallazgo se había puesto a aullar como si la hembra estuviese maltratándolo, llamando así a su propia madre, que en cuanto apareció alejó a la hembra.

Para que un animal mienta, fuera de toda instrucción genética, es necesario que perciba, con la experiencia, cuál es el efecto de su comportamiento en los otros. Parecen ser los primates los más hábiles a la hora de inventar sofisticadas mentiras, demostrando toda la flexibilidad, complejidad y variedad de engaños que les está permitida por sus desarrolladas capacidades cognitivas.

Los estudios con primates sugieren en estos un grado de flexibilidad cognitiva notable que les permite, entre otras cosas, recordar, predecir, planificar, anticipar e incluso comprender hasta cierto punto los elementos que constituyen sus problemas de forrajeo. Uno de los ingredientes fundamentales de dicha flexibilidad cognitiva es la habilidad para utilizar representaciones mentales. De esta manera, Call (2000) distingue tres tipos de representación mental en primates: la estática, la dinámica y la relacional. Las representaciones estáticas consisten en recordar ciertas características del entorno, por ejemplo, los monos ardilla (*Saimiri sciureus*), los capuchinos y los gorilas, entre otros, son capaces de recordar los lugares donde hallar alimento. Las representaciones dinámicas, por su parte, están dirigidas a la predicción de los cambios potenciales en el entorno. Dos monos rhesus, por ejemplo, son capaces de inferir los desplazamientos invisibles de un objeto en la pantalla de un ordenador. Por último, están las representaciones mentales relacionales, que están dirigidas a la codificación de las propiedades de los objetos en relación otros, por ejemplo, cuáles son las modificaciones necesarias para lograr convertir una rama con hojas en una herramienta adecuada para introducirla en un termitero. Los monos capuchinos, los gorilas y chimpancés y los macacos tienen una comprensión, aunque rudimentaria, de las relaciones de causa y efecto en problemas sencillos (Call, 2005).

En conclusión, es bastante variado el repertorio de ejemplos que indican que un considerable número de especies de animales no humanos, cuentan con un sistema cognitivo que les permite ciertas conductas de las que hace ya bastante tiempo se creían exclusivas del ser humano. En el caso de la intencionalidad, por ejemplo, “no tiene por qué ser una capacidad radical, en el sentido de que o se posee como la poseemos nosotros o, en caso contrario, no se posee. Puede sostenerse una visión gradualista de la misma” (Cuevas, 2016, 168).

Y es que parece innegable que muchas criaturas no lingüísticas se comportan de tal manera que merecen requerir el tratamiento de criaturas pensantes. Con las mascotas (mamíferos) que tenemos en casa, podemos percibir, por ejemplo, que algunos animales sienten dolor, alegría, tristeza y vergüenza. Básicamente, es cómodo atribuir pensamientos a otros animales por no contar con un modo mejor de explicar su comportamiento.

Sin embargo, a esto bien podría responder Davidson que “podemos seguir explicando la conducta de las criaturas que carecen de lenguaje atribuyéndoles actitudes proposicionales, aunque reconociendo al mismo tiempo que tales criaturas no poseen realmente actitudes proposicionales” (Davidson, 1982, 324).

Ya en líneas anteriores se afirmaba –atendiendo la sugerencia de Glock–, que el hecho de que los animales tengan o no conceptos de tipo simple, no depende de la posesión de lenguaje, sino de hasta qué punto sus discriminaciones están gobernadas por reglas y sean, de este modo, intencionales. La noción de seguir una regla es normativa en un sentido en que no lo son la noción de creencia y la de intención. Cuando un animal acata una regla actúa de tal manera que su actuar esté de acuerdo con una norma fijada independientemente de su actuación. En casos básicos en que la conducta de un animal puede ser descrita en términos de intenciones simples, sencillamente no hay tal norma (Prades, 2009). Cuando un perro, que no tiene lenguaje, trata de beber, no tiene por qué seguir una regla. Pero nuestra atribución de contenido sí es normativa; utilizamos el lenguaje y los conceptos para hacer la atribución, pero ello no demuestra que el animal siga una norma.

Como hemos visto hasta ahora, parece que es inevitable caer en el antropomorfismo cuando hablamos de algo tan complejo como los posibles procesos mentales en animales no humanos. ¿Cómo llamar, sino juego, a la acción de un chimpancé que simula golpear a otro más grande, con grandes vocalizaciones y dando pisotones al suelo en un contexto no agresivo? Así, se sugiere que la meta en estos casos no es erradicar el vocabulario antropomórfico de nuestros razonamientos, sino evitar caer en el antropomorfismo ingenuo, acrítico y dogmático. La antroponegación no es la alternativa cuando queremos evitar descripciones antropomórficas, sino la crítica racional y documentada.

La posición alternativa al antropomorfismo ingenuo y la antroponegación, es el antropomorfismo científico y documentado. Este, en lugar de centrarse en el punto de vista humano (“¿Cómo me sentiría yo en esa situación?”), está centrado en el animal. Es tomado más como un medio que como un fin y su objetivo no es encontrar una cualidad en el animal que sea equivalente a algún aspecto de nuestra propia vida interior. En lugar de ello, aprovechando que somos animales es posible desarrollar ideas, hipótesis que puedan probarse. El antropomorfismo científico sitúa a todos los animales, incluidos los humanos, en el mismo plano explicativo (De Waal, 2007). La aplicación del autoconocimiento humano para explicar el comportamiento en otros animales, más aún cuando están emparentados, es legítima. Al respecto, escribió Gordon Burghardt hace más de tres décadas:

Lo que pido es un antropomorfismo crítico, una inferencia predecible que estimule el uso de datos procedentes de diversas fuentes (experimentos previos, anécdotas, publicaciones, ideas y percepciones personales, el ponerse en el lugar del animal, observaciones naturalistas, etc.). Por



muy ecléctico que sea en su origen, el producto debe ser una inferencia capaz de ser demostrada o, por defecto, capaz de conducir a predicciones apoyadas en datos públicos (Burghardt, 1985, 917).

No tiene nada de anticientífico utilizar los mismos términos para referirse a seres humanos y animales; más aún cuando defendemos que un mismo fenómeno aparece en especies similares. La continuidad evolutiva sugiere un movimiento fluido de otros animales a seres humanos y de seres humanos a los demás animales. La idea no es buscar cualidades humanas en otros animales, sino comprender cómo son estos y utilizar el lenguaje y los conceptos que mejor se ajusten a lo que vemos. Lo llamemos como lo llamemos, debemos al menos estar de acuerdo en que los demás animales y los seres humanos parecen compartir muchas características, entre ellas las de tener emociones y pensamientos (Bekoff y Pierce, 2010). El antropomorfismo conceptual no es anticientífico.

Como bien afirma Cameron Buckner (2013), en un marco darwinista, no hay una buena razón para evitar conceptos simplemente porque se derivan de los comportamientos de nuestra especie. La aplicación de estos conceptos a los demás animales no solo enriquece la gama de hipótesis a considerar, sino que también cambia la visión de nosotros mismos: cuanto más parecidos a los humanos nos parecen otros animales, más como animales nos veremos a nosotros mismos.

## 5. CONCLUSIONES

El antropomorfismo científico parece ser una forma de parsimonia evolutivamente viable para los interesados en estudiar la conducta de los animales si es que el continuismo darwinista entre las personas y los demás animales ha de mantenerse como horizonte necesario de las ciencias de la conducta.

Pero, aunque compartamos la idea de la continuidad biológica o evolutiva con nuestros parientes evolutivos más cercanos y, por tanto, rechazemos la antroponegación, ¿existen algunas conductas o rasgos comportamentales únicos en el animal humano? La respuesta a esta pregunta poco podría importar si asumimos que somos únicos en la medida en que otras especies animales también lo son. La atribución de procesos mentales, intenciones, teoría de la mente o conceptos a otros animales sigue generando hoy intensos debates como décadas atrás, entre los que defienden irrestrictamente el antropomorfismo, quienes lo defienden con algunos matices, y quienes lo rechazan completamente. La etología

cognitiva -de Waal entre los etólogos más citados y criticados- continúa aportando material y argumentos al debate.

Una actividad tan frecuente, cotidiana y aparentemente inocente en la vida de los animales sociales como lo es el juego, por ejemplo, puede arrojar luces sobre por qué nos comportamos como lo hacemos y en qué medida los mecanismos sociales, no solamente determinan nuestra supervivencia, sino que la posibilitan.

## REFERENCIAS BIBLIOGRÁFICAS

- Allen, Colin (1999). Animal concepts revisited: the use of self-monitoring as an empirical approach. *Erkenntnis*, 51(1), 33-40. <https://doi.org/10.1023/A:1005545425672>
- Allen, Colin y Bekoff, Marc (1999). *Species of Mind. The Philosophy and Biology of Cognitive Ethology*. Cambridge: The MIT Press.
- Allen-Hermanson, Sean. (2005). Morgan's Canon Revisited. *Philosophy of Science*, 72(4), 608-631. <https://doi.org/10.1086/505187>
- Andrews, Kristin y Huss, Brian (2014). Anthropomorphism, anthropocentrism, and the null hypothesis. *Biol Philos*, (29), 711-729. <https://doi.org/10.1007/s10539-014-9442-2>
- Asensio, Norberto (2014). Etología social. En S. Sánchez López (Ed.), *Etología. La ciencia del comportamiento animal* (pp. 419-442). Barcelona: UOC.
- Bekoff, Marc y Pierce, Jessica (2010). *Justicia salvaje. La vida moral de los animales*. Madrid: Turner.
- Bermúdez, José Luis (2003). *Thinking without Words*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195159691.001.0001>
- Buckner, Cameron (2013). Morgan's Canon, meet Hume's Dictum: avoiding anthropofabulation in cross-species comparisons. *Biol Philos*, (28), 853-871. <https://doi.org/10.1007/s10539-013-9376-0>
- Burghardt, Gordon (1985). Animal awareness: Current perceptions and historical perspective. *American Psychologist*, (40), 905-919. <https://doi.org/10.1037/0003-066X.40.8.905>
- Byrne, Richard, y Corp, Nadia (2004). Neocortex size predicts deception rate in primates. *Proc. R. Soc. Lond. B*, (271), 1693-1699. <https://doi.org/10.1098/rspb.2004.2780>

- Call, Josep (2000). Representing space and objects in monkeys and apes. *Cognitive Science*, (24), 397-422. [https://doi.org/10.1207/s15516709cog2403\\_3](https://doi.org/10.1207/s15516709cog2403_3)
- Call, Josep (2001). Chimpanzee social cognition. *Trends in cognitive sciences*, 5(9), 388-393. [https://doi.org/10.1016/S1364-6613\(00\)01728-9](https://doi.org/10.1016/S1364-6613(00)01728-9)
- Call, Josep (2005). La evolución de la inteligencia ecológica en los primates: el problema de encontrar y procesar alimento. En F. Guillén-Salazar (Ed.), *Existo, luego pienso: los primates y la evolución de la inteligencia humana* (pp. 47-73). Madrid: Ateles Editores.
- Call, Josep (2006). Descartes' two errors: Reason and reflection in the great apes. En S. Hurley y M. Nudds (Eds.), *Rational Animals?* (pp. 219-234). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198528272.003.0010>
- Call, Josep (2010). Do apes know that they could be wrong? *Animal Cognition*, (13), 689-700. <https://doi.org/10.1007/s10071-010-0317-x>
- Castrodeza, Carlos (2009). *La darwinización del mundo*. Barcelona: Herder.
- Causey, Kayla y Bjorklund, David (2016). La evolución de la cognición. En V. Swami (Ed.), *Psicología evolucionista. Una introducción crítica* (pp. 42-83). México: Fondo de Cultura Económica.
- Colmenares, Fernando (2005). De laberintos sociales y de cómo salir de ellos: inteligencia social. En F. Guillén-Salazar (Ed.), *Existo, luego pienso: los primates y la evolución de la inteligencia humana* (pp. 75-128). Madrid: Ateles Editores.
- Colmenares, Fernando (2014). Etología social. En S. Sánchez López (Ed.), *Etología. La ciencia del comportamiento animal* (pp. 295-348). Barcelona: UOC.
- Crook, John y Goss-Custard, John (1972). Social Ethology. *Annual Review of Psychology*, 23, 277-312. <https://doi.org/10.1146/annurev.ps.23.020172.001425>
- Cuevas, Ana (2016). Artefactualidad animal. *Ludus Vitalis*, XXIV(45), 155-173.
- Darwin, Charles (1936)[1871]. *The Descent of Man and Selection in Relation to Sex*. New York: The Modern Library.
- Davidson, Donald (1982). Rational Animals. *Dialectica*, 36(4), 317-327. <https://doi.org/10.1111/j.1746-8361.1982.tb01546.x>
- Davidson, Donald (1999a). The emergence of Thought. *Erkenntnis*, 51 (1), 7-17.
- Davidson, Donald (1999b). Reply to Simon J. Evine. En L. Hanh (Ed.), *The Philisophy of Donald Davidson* (pp. 305-310). Chicago: Open Court.

- Davidson, Donald (2001). What Thought Requires? En J. Branquinho, *The Foundations of Cognitive Science* (pp. 121-132). New York: Oxford University Press. <https://doi.org/10.1093/oso/9780198238904.003.0008>
- Davidson, Donald (2003). Tres variedades de conocimiento. En D. Davidson, *Subjetivo, intersubjetivo, objetivo* (pp. 280-300). Madrid: Cátedra.
- De Waal, Frans (1989). *La política de los chimpancés*. Madrid: Alianza editorial.
- De Waal, Frans (1997). *Bien natural. Los orígenes del bien y del mal en los humanos y otros animales*. Barcelona: Herder.
- De Waal, Frans (2007). Seres moralmente evolucionados. En F. De Waal (Ed.), *Primates y filósofos. La evolución de la moral del simio al hombre* (pp. 23-111). Barcelona: Paidós.
- Delius, Juan y Godoy, Adrian (2001). Los procesos cognitivos en las aves y sus diversos orígenes. En J. Mora (Ed.), *Neuropsicología cognitiva: Algunos problemas actuales* (pp. 196-218). Aljibe: Archidona.
- Dennett, Daniel (2006). *Breaking the spell. Religion as a natural phenomenon*. New York: Penguin Books.
- Diéguez, Antonio (2005). ¿Hay diferencia esencial entre hombres y animales? Animales por derecho. *THÉMATA. Revista de Filosofía*, (35), 83-90.
- Diéguez, Antonio (2012). ¿Pueden pensar los animales? En J. Martínez y A. Ponce de León (Eds.), *Darwin y el evolucionismo contemporáneo* (pp. 313-326). México: Siglo XXI.
- Diéguez, Antonio (2014). Pensamiento conceptual en animales. En A. Diéguez y J. M. Atencia (Eds.), *Naturaleza animal y humana* (pp. 83-114). Madrid: Biblioteca Nueva.
- Dretske, Fred (1995). *Naturalizing the Mind*. Cambridge: The MIT Press.
- Dyer, Fred y Gould, James (1983). Honey bee navigation. *Amer. Sci.*, (71), 587-597.
- Fitzpatrick, Simon (2008). Doing Away with Morgan's Canon. *Mind & Language*, 23(2), 224-246. <https://doi.org/10.1111/j.1468-0017.2007.00338.x>
- Flombaum, Jonathan y Santos, Laurie (2005). Rhesus monkeys attribute perception to others. *Current Biology*, 15(5), 447-452. <https://doi.org/10.1016/j.cub.2004.12.076>
- Glock, Hans-Johann (2009). *La mente de los animales: problemas conceptuales*. Oviedo: KRK.
- Glock, Hans-Johann (2015). Comunicación personal vía correo electrónico con Hans-Johann Glock sobre procesos mentales en animales no humanos.

- Gould, James (1982). Why do honey bees have dialects. *Behav. Ecol. Sociobiol.*, (10), 53-56. <https://doi.org/10.1007/BF00296395>
- Gould, James (2002). Can honey bees create cognitive “maps”. En M. Bekoff, C. Allen y G. Burghardt (Eds.), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (pp. 41-46). Cambridge: MIT Press. <https://doi.org/10.7551/mitpress/1885.003.0009>
- Hare, Brian, Call, Josep y Tomasello, Michael (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, (61), 139-151. <https://doi.org/10.1006/anbe.2000.1518>
- Hare, Brian, Call, Josep y Tomasello, Michael (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, 101(3), 495-514. <https://doi.org/10.1016/j.cognition.2005.01.011>
- Herrnstein, Richard. (1990). Levels of stimulus control: a functional approach. *Cognition*, 37(1-2), 133-166. [https://doi.org/10.1016/0010-0277\(90\)90021-B](https://doi.org/10.1016/0010-0277(90)90021-B)
- Heyes, Cecilia (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, 21(1), 101-134. <https://doi.org/10.1017/S0140525X98000703>
- Jamieson, Dale y Bekoff, Marc (1993). On aims and methods of cognitive ethology. *Philosophy of Science Association*, (2), 110-124. <https://doi.org/10.1086/psaprocbienmeetp.1992.2.192828>
- Kaminski, Juliane, Call, Josep y Tomasello, Michael (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109(2), 224-234. <https://doi.org/10.1016/j.cognition.2008.08.010>
- Krupenye, Christopher, Kano, Fumihiro, Hirata, Satoshi, Call, Josep y Tomasello, Michael (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308), 110-114. <https://doi.org/10.1126/science.aaf8110>
- Llorente, Michael (2019). *Primates. Biología, comportamiento y evolución*. Barcelona: Lynx Edicions.
- Manning, Aubrey y Dawkins, Mariam S. (2012). *An Introduction to Animal Behaviour*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139030304>
- Mitchell, Sandra (2005). Anthropomorphism and cross-species modeling. En L. Daston y G. Mitman (Eds.), *Thinking with Animals: New Perspectives on Anthropomorphism* (pp. 100-117). New York: Columbia University Press.
- Morgado, Ignacio (2017). *Cómo percibimos el mundo. Una exploración de la mente y los sentidos*. Bogotá: Ariel.
- Morgan, C. Lloyd (1896). *Habit and Instinct*. London: Edward Arnold. <https://doi.org/10.1037/12922-000>

- Morgan, C. Lloyd (1903). *An Introduction to Comparative Psychology*. Londres: Scott. <https://doi.org/10.1037/13701-000>
- Mulcahy, Nicholas y Call, Josep (2006). Apes save tools for future use. *Science*, (312), 1038-1040. <https://doi.org/10.1126/science.1125456>
- Newen, Albert y Bartels, Andreas (2003). Animal Minds and the Possession of Concepts. *Philosophical Psychology*, 20(3), 283–308. <https://doi.org/10.1080/09515080701358096>
- Osvath, Mathias y Karvonen, Elin (2012). Spontaneous Innovation for Future Deception in a Male Chimpanzee. *PLoS ONE*, 7(5), e36782. <https://doi.org/10.1371/journal.pone.0036782>
- Penn, Derek y Povinelli, Daniel (2013). The Comparative Delusion: the 'behavioristic' / mentalistic' dichotomy in comparative Theory of Mind research. En J. Metcalfe y H. Terrace (Eds.), *Agency and Joint Attention* (pp. 62-80). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199988341.003.0004>
- Povinelli, Daniel, Bering, Jesse y Giambrone, Steve (2000). Toward a science of other minds: escaping the argument by analogy. *Cognitive Science*, 24(3), 509-541. [https://doi.org/10.1207/s15516709cog2403\\_7](https://doi.org/10.1207/s15516709cog2403_7)
- Prades, Josep (2009). Introducción. Expresión y contenido intencional. En H.-J. Glock (Ed.), *La mente de los animales: problemas conceptuales* (pp. 11-56). Oviedo: KRK.
- Pruetz, Jill y Bertolani, Paco (2007). Savanna Chimpanzees, Pan troglodytes verus, Hunt with Tools. *Current Biology*, 17(5), 412-417. <https://doi.org/10.1016/j.cub.2006.12.042>
- Quinn, Paul, Adams, Adria, Kennedy, Erin, Shettler, Lauren y Wasnik, Amanda (2003). Development of an abstract category representation for the spatial relation between in 6- to 10-month-old infants. *Developmental Psychology*, 39(1), 151-163. <https://doi.org/10.1037/0012-1649.39.1.151>
- Quinn, Paul y Tanaka, James (2007). Early development of perceptual expertise: Within-basic-level categorization experience facilitates the formation of subordinate-level category representations in 6- to 7-month-old infants. *Memory & Cognition*, 35(6), 1422-1431. <https://doi.org/10.3758/BF03193612>
- Rollin, Bernard (1997). Anecdote, Anthropomorphism, and Animal Behavior. En R. Mitchell, N. Thompson y L. Miles (Eds.), *Anthropomorphism, Anecdotes, and Animals* (pp. 125-133). New York: State University of New York Press.
- Shettleworth, Sara (2010). *Cognition, communication, and behavior*. New York: Oxford. <https://doi.org/10.1093/oso/9780195319842.001.0001>

- Sober, Elliott (2005). Comparative Psychology Meets Evolutionary Biology. Morgan's Canon and Cladistic Parsimony. En L. Daston y G. Mitman (Eds.), *Thinking with animals: New perspectives on anthropomorphism* (pp. 85-99). New York: Columbia University Press.
- Spelke, Elizabeth (1990). Principles of Object Perception. *Cognitive Science*, (14), 29-56. [https://doi.org/10.1207/s15516709cog1401\\_3](https://doi.org/10.1207/s15516709cog1401_3)
- Spelke, Elizabeth (1994). Initial knowledge: six suggestions. *Cognition*, (50), 431-445. [https://doi.org/10.1016/0010-0277\(94\)90039-6](https://doi.org/10.1016/0010-0277(94)90039-6)
- Spelke, Elizabeth (1998). Nativism, empiricism, and the origins of Knowledge. *Infant Behavior and Development*, 21(2), 181-200. [https://doi.org/10.1016/S0163-6383\(98\)90002-9](https://doi.org/10.1016/S0163-6383(98)90002-9)
- Spelke, Elizabeth y Tsivkin, Sanna (2003). Initial knowledge and conceptual change: space and number. En M. Bowerman y S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 70-97). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620669.005>
- Stern, Daniel (1991). *El mundo interpersonal del infante: una perspectiva desde el psicoanálisis y la psicología evolutiva*. Barcelona: Paidós.
- Thompson, Roger y Oden, David (2000). Categorical perception and conceptual judgements by nonhuman primates: the paleological monkey and the analogical ape. *Cognitive Science*, 24(3), 363-396. [https://doi.org/10.1207/s15516709cog2403\\_2](https://doi.org/10.1207/s15516709cog2403_2)
- Tomasello, Michael (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press. <https://doi.org/10.4159/9780674044371>
- Tomasello, Michael, Call, Josep y Hare, Brian (2003). Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends in Cognitive Science*, 7(4), 153-156. [https://doi.org/10.1016/S1364-6613\(03\)00035-4](https://doi.org/10.1016/S1364-6613(03)00035-4)
- Vauclair, Jacques (2002). Categorization and conceptual behavior in nonhuman primates. En M. Bekoff, C. Allen y G. Burghardt (Eds.), *The Cognitive animal. Empirical and theoretical perspectives on animal cognition* (pp. 239-245). Cambridge: The MIT Press. <https://doi.org/10.7551/mitpress/1885.003.0034>
- Visalberghi, Elisabetta, y Fragaszy, Dorothy (2013). The Etho-Cebus Project: Stone-tool use by wild capuchin monkeys. En C. Sanz, J. Call y C. Boesch (Eds.), *Tool Use in Animals: Cognition and Ecology* (pp. 203-222). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511894800.013>
- Vonk, Jennifer y MacDonald, Suzanne (2002). Natural concepts in a juvenile gorilla (gorilla gorilla gorilla) at three levels of abstraction. *Journal*

- of the Experimental Analysis Behavior*, 78(3), 315-332. <https://doi.org/10.1901/jeab.2002.78-315>
- Vonk, Jennifer y MacDonald, Suzanne (2004). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology*, 118(1), 3-13. <https://doi.org/10.1037/0735-7036.118.1.3>
- Wright, Robert (2007). Los usos del antropomorfismo. En F. de Waal (Ed.), *Primates y filósofos. La evolución de la moral del simio al hombre* (pp. 115-130). Barcelona: Paidós.
- Wuketits, Franz (1984). Gnoseología evolutiva: el nuevo desafío. En K. Lorenz y F. Wuketits (Eds.), *La evolución del pensamiento* (pp. 11-27). Barcelona: Argos Vergara.



eISSN: 1989-3612

DOI: <https://doi.org/10.14201/art2024.30762>

# HOMOLOGÍA Y ANTROPOMORFIZACIÓN CRÍTICA

## *Homology and Critical Anthropomorphization*

Julio TORRES MELÉNDEZ 

*Departamento de Filosofía, Universidad de Concepción, Chile*  
jutorres@udec.cl

Claudia MUÑOZ TOBAR 

*Departamento de Filosofía, Universidad de Concepción, Chile*  
claumuno@udec.cl

Recibido: 26/10/2022

Revisado: 02/05/2023

Aceptado: 28/06/2023

**RESUMEN:** A partir de un examen crítico de la distinción predarwiniana entre actividad y uso con la que Alan Love (2007) intenta resolver la suelta contradicción contenida en el concepto de homología de función, argumentamos que es prescindible la estrategia de antropomorfización defendida por Frans de Waal en el estudio de la conducta social de primates. Lo que se requiere para eludir el compromiso con la antropomorfización es hacer explícito el criterio de identidad del rasgo biológico que está contenido en el concepto de homología filogenética. Se requiere asimismo revelar la estructura jerárquica y desconectada de las homologías de categorías psicológicas. Sostenemos que hacer explícita esa estructura jerárquica y la posibilidad de realizabilidad múltiple de patrones de comportamiento social que cumplen funciones homólogas, le permitiría a de

Waal responder de manera efectiva a los críticos que niegan sus tesis de la continuidad evolutiva de la moral.

*Palabras clave:* de Wall, función biológica, identidad de rasgo, filogenia, realización múltiple.

**ABSTRACT:** Based on a critical examination of the Pre-Darwinian distinction between activity and use with which Alan Love (2007) attempts to resolve the alleged contradiction contained in the concept of homology of function, we argue that the strategy of anthropomorphization advocated by Frans de Waal in the study of primate social behavior is dispensable. What is required to circumvent the commitment to anthropomorphization is to make explicit the criterion of biological trait identity that is contained in the concept of phylogenetic homology. It is also required to reveal the hierarchical and disconnected structure of psychological category homologies. We argue that making explicit that hierarchical structure and the possibility of multiple realizability of patterns of social behavior that fulfill homologous functions would enable de Waal to respond effectively to critics who deny his thesis of the evolutionary continuity of morality.

*Keywords:* de Wall, homology, biological function, trait identity, phylogeny; multiple realizability.

## 1. INTRODUCCIÓN

La argumentación en defensa de la excepcionalidad humana es un tópico en la historia del pensamiento filosófico. No se han dejado dimensiones de la vida humana sin escudriñar en búsqueda de lo que nos haría esencialmente singulares. La experiencia consciente, la autoconciencia, el lenguaje, la voluntad, la racionalidad, el libre albedrío, la vida moral, la vida social y la organización política, pueden, se admite, no ser todas ellas propiedades o capacidades exclusivas de la forma de vida humana, pero sus modos de darse en nosotros y la manera en que se articulan entre sí estas cualidades o facultades para producir la acción y la cognición humanas, nos hacen únicos y nos dan razones para satisfacer nuestra permanente ansia de singularidad. Esta búsqueda de la excepcionalidad humana no ha sido del todo infructuosa, pero ha sido sin duda sesgada. Se ha centrado solo en las diferencias, como si no hubiera nada que aprender acerca de lo que Darwin llamó las 'verdaderas afinidades' entre las distintas especies animales. El pensamiento evolucionista y en particular la teoría darwiniana del ancestro común universal de la vida, no parecen haber alterado esta búsqueda de la excepcionalidad de

manera significativa. Aunque es muy improbable que hoy se defienda que todos los animales sean autómatas que carecen de genuinas experiencias conscientes o que sean solo cosas con las que “se puede hacer y deshacer a capricho”<sup>1</sup>, seguimos percibiéndonos como seres singulares, seres sin historia evolutiva<sup>2</sup>. Frans de Waal ha intentado modificar no sin resistencias esta trayectoria del pensamiento filosófico en sus estudios acerca del comportamiento de primates no humanos en el marco de una investigación acerca de la evolución cognitiva (de Waal, 2021, 42). Es en este contexto que debe entenderse, por ejemplo, su interés por la teoría de la capa, un caso de la tesis de la singularidad de acuerdo con la cual la moralidad humana es un barniz cultural generado para cubrir y también contrarrestar una naturaleza humana amorala, egoísta y violenta. La tesis de la continuidad evolutiva supone, en cambio, que la moralidad se asienta sobre capacidades cognitivas que se han desarrollado a partir de un núcleo emocional básico e indispensable de origen mamífero: la empatía. De este mismo núcleo depende también la existencia de una vida mental e intencional en animales sociales. Este núcleo es la base sobre la que se han construido las capacidades cognitivas que consideramos propiamente humanas, como la atribución intencional en niveles que superan la teoría de la mente, la adopción plena de la perspectiva ajena, la cooperación o la moralidad misma. De esta manera, la cognición humana es, para de Waal, una variante de la cognición animal (de Waal, 2021, 17).

Es en el contexto de la tesis de la continuidad evolutiva en donde los conceptos de antropomorfismo y de semejanza homológica resultan centrales. De ahí que pretendemos, por una parte, hacer un examen crítico del uso heurístico de la antropomorfización defendido por de Waal en el estudio de la conducta social de primates y, por otra parte, mostrar la centralidad que debería tener ahí la homología de función. Específicamente, abordaremos críticamente ese enfoque antropomórfico a partir de un examen crítico de la distinción predarwiniana entre *actividad* y *uso* con la que Alan Love (2007) intenta resolver la supuesta contradicción contenida en el concepto de homología de función. Sostendremos que

1. Kant, 1991 [1800], 15-16.

2. Las ideas cartesianas acerca de la naturaleza de la experiencia animal no han desaparecido tan pronto como lo pudiéramos pensar hoy. Frans de Waal nos recuerda el extendido desprecio hacia el sufrimiento animal de la comunidad científica de la década de 1970: “era entonces dominante el punto de vista que sostenía que los animales no eran sino meros robots, desprovistos de sentimientos, ideas o emociones” (de Waal, 2007, 106).

la homología de función es defendible sin abandonar la noción filosófica de la función como el efecto seleccionado. Lo que se requiere para eludir el compromiso con la antropomorfización es hacer explícito el criterio de identidad del rasgo biológico que está contenido en el concepto de homología filogenética. Se requiere asimismo mostrar la estructura jerárquica y desconectada, o de realización múltiple, de las homologías de comportamiento (Ereshefsky, 2012). Esta estructura la encontramos en los elementos constituyentes de la empatía animal. De Waal (2014) ha identificado esta estructura como un conjunto de capas organizadas jerárquicamente en torno a un núcleo compartido por diversas especies animales. A partir de la capa interna, que de Waal denomina 'correspondencia de estados' (también la llama 'empatía corporal' y 'sincronía corporal'), se desarrollan otras dos, la preocupación por los demás y la adopción de la perspectiva ajena. Argumentamos que los rasgos de comportamiento asociados a estas capas, como el contagio emocional, la consolación y la ayuda orientada, constituyen una jerarquía de homologías de función de realización múltiple.

## 2. ANTROPOMORFISMO Y HOMOLOGÍA DE FUNCIÓN

A pesar de la importancia que tiene la identificación de homologías de rasgos funcionales para el estudio comparativo de expresiones de estados psicológicos, el mismo concepto de homología de una función ha resultado problemático dada la oposición entre explicación homológica y explicación funcional o adaptativa. De acuerdo con la definición darwiniana de la homología como identidad por origen común, donde hay explicación homológica no puede haber explicación funcional o adaptativa y viceversa. La semejanza analógica se explica como el resultado de procesos semejantes de adaptación. Si dos estructuras o dos patrones de comportamiento han evolucionado de manera independiente para cumplir funciones semejantes entonces la semejanza de estas funciones es analógica y no homológica. Si, por ejemplo, un primatólogo pretende afirmar que hay evidencia suficiente para sostener que ciertas expresiones faciales realizan, en determinada familia de primates, una función *semejante* a la risa humana, entonces se esperaría que hubiera claridad acerca del concepto de homología de una función para establecer de qué naturaleza es esta semejanza, pero este no parece ser el caso. De ahí que Love (2007) nos advierta que hay cierta resistencia filosófica a explicar determinadas semejanzas funcionales entre distintas especies animales sobre la base de una hipótesis de homología darwiniana, debido a la supuesta

oposición entre homología y patrón funcional. Así, por ejemplo, Schmidt y Cohn en un estudio evolutivo de las expresiones faciales humanas sostienen que:

Los enfoques evolutivos [...] se basan necesariamente en el estudio de las expresiones faciales como adaptaciones biológicas, aunque gran parte de los trabajos recientes sobre la expresión facial se han realizado fuera de esta perspectiva. No obstante, los resultados de este tipo de estudio funcional de la expresión facial proporcionan una base sobre la que fundamentar las hipótesis evolutivas del comportamiento facial (Schmidt y Cohn, 2001, 20).

Este puede ser un buen ejemplo que muestra la renuencia a aplicar un enfoque homológico en los estudios funcionales de expresiones de estados internos, aunque esto es justamente lo que requieren los estudios comparativos de este tipo. No hay tal *necesidad* de que los enfoques evolutivos traten a las expresiones faciales como adaptaciones, de hecho, el mismo Darwin le negaría a gran parte de ellas el carácter adaptativo en su estudio pionero al respecto (Darwin, 2009 [1890])<sup>3</sup>. En la discusión contemporánea esta resistencia a la homología de función se debería, según Love, a la extendida idea filosófica de la función como el efecto seleccionado (la concepción etiológica de la función). De acuerdo con este punto de vista, la función de una estructura biológica o de un patrón de comportamiento, es el efecto de esa estructura o de ese comportamiento, en donde reside la explicación de por qué ha sido seleccionado y propagado en la historia de la especie. Si la explicación de la naturaleza de la función biológica reside en la selección natural, parece ser entonces una contradicción sostener que existen homologías de función, dado que la explicación homológica excluye justamente la explicación por selección. De ahí que Love afirme que:

La homología funcional es [...] una especie de error categorial, ya que lo que hace una estructura no debería ser objeto de una evaluación homológica y la similitud de función no es generalmente el resultado de un origen común, sino de la adaptación por selección natural a demandas ambientales comunes. La yuxtaposición de 'función' y 'homología' es oximorónica, o más bien una contradicción en términos teóricos (Love, 2007, 693).

Esta supuesta contradicción podría quizás explicar por qué de Waal cuando requiere justificar teóricamente la posibilidad de la comprensión de porciones significativas del comportamiento animal a semejanza de la

3. Cf. Fridlund, 1984 y Ginnobili, 2022.

comprensión de la acción humana, se decide por justificar una variante crítica del antropomorfismo y no la validez teórica de la homología de función. Parte de esta justificación se apoya en un paralelismo entre, por un lado, la semejanza metafórica que se pone en juego, muchas veces, en la ciencia y, por otro, la semejanza antropomórfica que, singularmente, requeriría la primatología (cf. de Waal, 1999, 264). De ahí que el antropomorfismo sea para de Waal un recurso heurístico para proponer hipótesis de homología funcional: "El objetivo final del científico que antropomorfiza no es la proyección más satisfactoria de los sentimientos humanos en el animal, sino obtener ideas contrastables y observaciones replicables" (de Waal, 2006, 63). No se trataría, nos dice, de un recurso antropocéntrico, sino más bien animal-céntrico, en la medida que exige una profunda familiaridad con el comportamiento del animal y la historia natural de la especie. De acuerdo con de Waal "cuanto más cercana a nosotros sea una especie, más contribuirá el antropomorfismo a su comprensión y mayor será el perjuicio de la antroponegación. Y viceversa, cuanto mayor sea la distancia entre una especie y nosotros, mayor será el riesgo de que el antropomorfismo sugiera similitudes que tienen un origen independiente" (de Waal, 2021, 38). Para el primatólogo, podría argumentarse, la estrategia de antropomorfización, en tanto recurso heurístico para la formulación de "ideas contrastables y observaciones replicables" y, finalmente, para el descubrimiento de homologías de función, es simétrica a la estrategia que puede utilizar el biólogo evolucionista para descubrir qué estructuras o patrones de comportamiento son adaptaciones para una determinada especie. Aquí el investigador empírico puede usar un enfoque intencional, de ingeniería inversa, como recurso heurístico para hacer hipótesis acerca de qué efectos del rasgo constituyen una adaptación a las demandas del medio. El evolucionista, para descubrir su función, "trata de figurarse" la *razón* por la cual la naturaleza hizo un rasgo tal como lo hizo (Dennett, 1995, 213). Hay, sin embargo, una diferencia fundamental entre ambas estrategias. Luego de la explicación por selección la explicación intencional se autoelimina, no forma parte de la explicación evolucionista, pero no ocurre así con la explicación antropomórfica del primatólogo que descubre homologías y excluye similitudes que tienen un origen independiente. La atribución antropomórfica de determinadas representaciones mentales a primates no humanos no queda eliminada luego de la propuesta de explicación, pues esta atribución constituye por sí misma la explicación de la acción animal: si el primate no humano tiene un determinado comportamiento semejante a nosotros entonces asumimos que dada esta semejanza antropomórfica su comportamiento se *explica* porque es causado por semejantes representaciones mentales.

La explicación es antropomórfica, pero no ocurre así con la estrategia evolucionista de la ingeniería inversa: la explicación no es intencional, es evolucionista. Es por ello que justificar este enfoque antropomórfico como un recurso heurístico es anómalo: no representa correctamente el programa de investigación que lleva adelante de Wall. Y, por esto, es también prescindible, dada la larga tradición científica, que se remonta al pensamiento formalista pre-darwiniano, que puede explicar esa *semejanza antropomórfica* sobre la base de hipótesis de semejanza homológica con sólidas evidencias empíricas no antropomórficas (Owen, 1843; Owen, 2012 [1849]). La semejanza homológica, a diferencia de la semejanza antropomórfica, nunca es un recurso heurístico, sino que siempre es por sí misma una hipótesis contrastable. Lo que requiere el primatólogo no es entonces un enfoque de semejanza antropomórfica, sino una teoría de la representación mental que explique la conducta animal sobre la base de que sus representaciones cumplen funciones biológicas semejantes a las representaciones humanas, es decir, una teoría de la identidad funcional de la representación mental<sup>4</sup>.

### 3. ¿ES NECESARIO ABANDONAR LA CONCEPCIÓN ETIOLÓGICA DE LA FUNCIÓN?

Love sostiene que puede resolverse la aparente contradicción entre explicación homológica y explicación funcional recurriendo a la distinción pre-darwiniana entre *uso* y *actividad*. Esta distinción permitiría dar un significado apropiado al concepto de homología de función. El *uso* funcional de un rasgo biológico representa “para qué sirve”, mientras que la *actividad* funcional de un rasgo biológico captura “lo que hace” (o cómo funciona). Un determinado uso puede estar instanciado en diversas actividades y una actividad puede tener múltiples usos. Los términos no tienen

4. Esta teoría se ha formulado, se trata de la teoría teleosemántica que David Papineau y otros han contribuido a desarrollar sobre la base de la interpretación etiológica de la función biológica. La teleosemántica no requiere de un compromiso con un enfoque antropomórfico, pues se trata de una teoría unificada de las representaciones mentales animales y humanas. Asume que ellas tienen una misma explicación en la teoría darwiniana. De acuerdo con Papineau debemos entender la representación mental como un fenómeno biológico: “las representaciones son estados cuya función biológica es guiar el comportamiento de forma adecuada en determinadas condiciones” (Papineau, 2017, 97; cf. Syfarth *et al.*, 1980, 802). La teleosemántica, sin embargo, tiende a olvidar el componente histórico de la explicación evolucionista y, por esta razón, malentendiendo la naturaleza del criterio evolucionista de identidad de rasgo (cf. Torres, 2021).

equivalencia extensional. “Tampoco son equivalentes desde el punto de vista epistémico, ya que podemos saber que se hace algo, pero no para qué sirve y, alternativamente, podemos saber para qué sirve algo pero no cómo lo consigue” (Love 2007, 695-696). De ahí que sostenga Love que:

[...] el significado apropiado de ‘homología de función’ es el de actividad, no el de rol causal, ya que la actividad puede permanecer constante ‘bajo cualquier variedad de forma y función’. La evaluación de la similitud debida a la ascendencia común se hace por separado del rol que desempeña la función, ya sea entendido en términos de un rol causal, una ventaja en aptitud o una historia de selección. Necesitamos una interpretación de la función en términos de ‘qué es’ y no de ‘para qué sirve’, y la función como actividad se ajusta a este requerimiento (Love, 2007, 695).

De Waal nos ofrece un buen ejemplo en donde podemos aplicar esta distinción. El patrón de comportamiento *cara con los dientes descubiertos* (una *actividad* funcional), constituye una expresión típica entre los primates que consiste en “la exhibición facial en donde los primates muestran sus dientes y encías mediante la retracción de los labios” (de Waal, 1999, 267). De acuerdo con de Waal, esta expresión puede tener distinto significado o una “distinta interpretación funcional” (distintos *usos* funcionales) (Schmidt y Cohn, 2001, 20). En los chimpancés puede significar miedo aunque también es observable en contextos amistosos. En la *Política de los chimpancés* de Waal llama ‘sonrisa’ a esta expresión en tales contextos amistosos<sup>5</sup>. En el bonobo parece predominar esto último como también ocurre en los humanos. En un estudio comparativo del comportamiento comunicativo de bonobos y chimpancés de Waal había hecho un registro más preciso al respecto:

El rostro de dientes descubiertos y silente del chimpancé es idéntico al del bonobo. Se utiliza en la misma gama de circunstancias, con un acento en las funciones de apaciguamiento y consuelo. Las llamadas “sonrisas de placer” son menos comunes o menos conspicuas en el chimpancé que en el bonobo (de Waal, 1988, 194).

Estudios posteriores mostraron la consistencia entre el patrón acústico de las vocalizaciones de risa en distintos primates y los estudios

5. En la versión al español de *Chimpanzee Politics* se traduce ‘*grin*’ [sonrisa] por ‘mueca parecida a una sonrisa’ (por ejemplo, en de Waal, 2022, 53, 185-186), es decir, se hace el tipo de interpretación conservadora en términos de pautas conductuales (y no de expresión de emociones) que el mismo de Waal anuncia que no hará en su libro (de Waal, 2022, 73).



genéticos y morfológicos previos que establecieron sus relaciones filogenéticas. De acuerdo con Davila-Ross *et al.* (2010) los resultados de este estudio proporcionan evidencias de la ascendencia común de la risa en los seres humanos y de las vocalizaciones inducidas por las cosquillas en los grandes simios. Estos resultados:

apoyan la afirmación más general de la continuidad filogenética de las expresiones emocionales desde las manifestaciones no humanas hasta las humanas. La "risa", por tanto, no es un término antropomórfico y, en cambio, se puede rastrear como un tipo de vocalización hasta, al menos, el último ancestro común de los grandes simios modernos y los seres humanos, hace aproximadamente diez a dieciséis millones de años (Davila-Ross *et al.*, 2010, 192).

Tal como lo advierte de Waal los comportamientos *cara con los dientes descubiertos* son expresiones faciales homólogas para chimpancés, bonobos y humanos. Como también ocurre con el patrón acústico de las vocalizaciones de risa de estos primates. Pero se trataría en ambos casos, de acuerdo con la interpretación de Love, de una homología de la función entendida como una homología de la actividad y no del uso.

Pero, ¿por qué no habría homología de la función de uso? En ese mismo ejemplo puede asumirse, hipotéticamente, que el uso que los bonobos, chimpancés y humanos hacen del comportamiento *cara con los dientes descubiertos* es el mismo y que ese uso evolucionó solo una vez en el ancestro común de bonobos y humanos. Love no niega que exista homología de función de uso, pero se muestra escéptico acerca de las posibilidades de identificarlas dada la inestabilidad de estas homologías. Lo ejemplifica con el caso de la función de actividad bioquímica de los genes que se ha demostrado tiende a conservarse incluso en especies distantes filogenéticamente. Pero con el uso, o el papel causal, de esta función bioquímica, en los procesos de desarrollo ontogénico en distintas especies, no ocurre lo mismo: "Los mismos genes reguladores que son evolutivamente estables en términos de actividad-función son evolutivamente lábiles en términos de uso-función" (Love, 2007, 697). Esto hace que esas homologías de la función de uso sean "más sospechosas" o "problemáticas". De ahí que sostenga, como se ha dicho, que "el significado apropiado de 'homología de función' es el de actividad". Pero el problema, como admite, no es conceptual. No hay nada en el significado de 'homología de función de uso' que sea dudoso, la cuestión es meramente epistémica. Tiene que ver con la posibilidad empírica de determinar cuándo hay este tipo de homología. Esto, por cierto, muestra el carácter convencional de la distinción entre actividad y uso cuando ella

es utilizada como argumento para excluir una jerarquía de homología de función que al primatólogo le resulta fundamental en la identificación de patrones conductuales que, en distintas especies, cumplen los mismos roles causales en la expresión de estados internos. La posibilidad de que tales roles causales sean cumplidos por distintas funciones de actividad hace asimismo que la homología de actividad sea, en contra de lo que sostiene Love, menos relevante en este contexto. No parece que la función de actividad tenga la estabilidad que requiere Love. El mismo tipo de ejemplo en el nivel de la organización genética muestra esto, dado que "las estructuras homólogas pueden desarrollarse a partir de genes no homólogos y los genes homólogos pueden estar implicados en la producción de estructuras no homólogas en diferentes especies" (Brigandt, 2007, 720). Los mecanismos genéticos y celulares de desarrollo de las extremidades pueden variar, no ser homólogos, por ejemplo, en anuros respecto de otros vertebrados, pero aun así no hay duda de que esas extremidades deben ser consideradas homólogas<sup>6</sup>. En contraste con esto, los genes que dirigen el desarrollo de los ojos en los vertebrados y en los invertebrados pueden ser homólogos a pesar de que estas estructuras no lo sean (Hall, 1999, 345). Esta complejidad muestra que la jerarquía de niveles biológicos, entre ellos, moléculas, genes, tejidos, módulos de desarrollo ontogénico, órganos, poblaciones, patrones de conducta y categorías psicológicas, aunque estén causalmente conectados pueden estar homológicamente desconectados, es decir, pueden constituir lo que Marc Ereshefsky llamó 'el fenómeno de la jerarquía desconectada' de los patrones de homología (Ereshefsky, 2012; Hall, 1999). Este fenómeno resulta fundamental para comprender la relación que tienen las categorías psicológicas homólogas con los niveles biológicos inferiores. De acuerdo con Ereshefsky:

Si una categoría psicológica es una homología, entonces podemos tener una explicación de la realizabilidad múltiple de esa categoría. Las instancias de la misma categoría psicológica pueden variar en sus subestructuras físicas. También pueden variar en sus causas de desarrollo.

6. De acuerdo con Brian Hall, es "más útil distinguir los procesos de desarrollo como causas de las estructuras homólogas que considerarlos como la homología en sí misma. El desarrollo es un proceso, la homología es un patrón. Obviamente, es importante encontrar las 'reglas de construcción' para el desarrollo y la evolución de las extremidades [...], pero una extremidad construida sobre un conjunto de reglas no pierde su homología con las extremidades construidas sobre reglas diferentes; simplemente puede haber habido más divergencia y evolución de los procesos de desarrollo entre los dos grupos que dentro de cada grupo. Después de todo, el cambio es la esencia de la evolución" (Hall, 1999, 347).

Sin embargo, son instancias de la misma categoría debido a su filogenia compartida. [...] la filogenia proporciona las condiciones de identidad de una categoría psicológica, y la ontogenia explica la variación estable entre sus instancias (Ereshefsky, 2007, 667).

El ejemplo de Ereshefsky es el miedo no-doxástico que, a pesar de las diferencias subyacentes en el soporte biológico y en los contextos sociales o culturales en donde se expresa, tanto en humanos como en otros primates, constituye una misma categoría psicológica debido a la filogenia compartida. Debe advertirse que las relaciones de semejanza entre caracteres constituyen solo una evidencia hipotética de que hay homología. El criterio definidor lo constituye la filogenia compartida (Ghiselin, 2005; Roffé, 2020). La homología no es entonces una relación de semejanza cualitativa, sino de mismidad o de "identidad de estructura" tal como lo sostuvo Darwin (2009 [1890], 381; citado por Ginnobili, 2022, 37).

#### **4. LA SOLUCIÓN DE DARWIN A LA CONTRADICCIÓN CONTENIDA EN LA HOMOLOGÍA DE FUNCIÓN**

Hay otra manera de resolver lo que parece ser una contradicción entre explicación homológica y explicación funcional que no requiere abandonar la concepción etiológica de la función ni tampoco retroceder a la distinción predarwiniana entre uso y actividad. Darwin mismo abordó esta supuesta contradicción en un pasaje del *Origen* que no ha recibido suficiente atención y la identificó como "la *aparente paradoja* de que exactamente los mismos caracteres sean analógicos cuando una clase u orden es comparada con otra, pero que den verdaderas afinidades cuando los miembros de la misma clase u orden son comparados uno con otro" (Darwin, 1859, 427; énfasis añadido). Se expresa aquí la posibilidad de la coincidencia extensional entre el conjunto de caracteres homológicos y el conjunto de caracteres analógicos<sup>7</sup>.

7. Se trata de una posibilidad que tiene excepciones. En el caso de los rasgos vestigiales o rudimentarios no parece posible un contexto contrastivo en donde estos rasgos, bajo el supuesto de que no cumplen ninguna función, puedan ser interpretados como rasgos analógicos. De ahí la importancia que los taxónomos pre-evolucionistas le atribuyen en la clasificación. Darwin reinterpreta esta práctica a la luz de su propia teoría de la clasificación genealógica. "La comunidad de descendencia es el vínculo oculto que los naturalistas han estado buscando inconscientemente" (Darwin, 1859, 420).

Lo que debemos comprender es que la explicación homológica, al igual que la explicación por selección, tiene una naturaleza contrastiva<sup>8</sup>. La propiedad de ser una homología o ser una analogía depende del contexto de contraste. No advertir esto puede ser causa de confusiones incluso para quienes este concepto resulta ser especialmente central:

Para un cladista, las plumas son una sinapomorfía, o una homología, de las aves, como si las plumas mismas constituyeran, comprendieran o compusieran una homología particular que forma parte de la materia que constituye, comprende o compone las aves. Sin embargo, debe quedar claro que la homología, en general y en particular, es una relación entre homólogos (por ejemplo, plumas de diferentes aves) y no simplemente los homólogos en sí mismos (G. Nelson, en Hall, 1994, 120).

La forma tubular de la ballena y la forma de aleta de sus miembros, dice Darwin, es una semejanza analógica respecto de la forma tubular y las aletas de un pez, pues estas formas evolucionaron de manera independiente para cumplir la misma función en el ambiente marino. Pero hay semejanza homológica cuando se comparan estas formas con los diversos miembros del género o la familia al que pertenecen las ballenas: “entre los diferentes miembros de la familia de las ballenas, la forma del cuerpo y los miembros en forma de aleta presentan caracteres que ponen de manifiesto afinidades en toda la familia, no podemos dudar de que han sido heredadas de un antepasado común” (Darwin, 2010 [1872], 577; cf. Darwin, 1859, 428). Lo que muestra el ejemplo de Darwin es que la presencia causal del ancestro común puede excluir la hipótesis del efecto causal de la selección y, a su vez, que la exclusión de la hipótesis del ancestro común puede mostrar la acción causal de la selección natural. Si aplicamos la idea darwiniana de la posibilidad de la coincidencia extensional entre homología y analogía, entonces el comportamiento *cara con los dientes descubiertos* de los chimpancés, bonobos y humanos constituyen una semejanza homológica para estas especies de la familia de los primates, pero constituye una analogía, asumamos, si la comparamos con el comportamiento *cara con los dientes descubiertos* del perro doméstico. Presumiblemente estos comportamientos evolucionaron de manera independiente en estos dos grupos. Pero la función de este comportamiento pudo haberse originado solo una vez en el ancestro que

8. Acerca de la naturaleza contrastiva de la explicación por selección natural, cf. Razeto-Barry y Frick, 2011.

comparten humanos, bonobos y chimpancés<sup>9</sup>. En este caso, la sonrisa de humanos, bonobos y chimpancés es una analogía respecto de comportamientos semejantes en otros mamíferos dado su rol causal en determinados entornos sociales, pero ese mismo comportamiento es una homología al interior del grupo que forman estas tres especies tal como lo muestran estudios filogenéticos basados en la acústica de las vocalizaciones de risa inducidas en orangutanes, gorilas, chimpancés, bonobos y humanos (Davila-Ross *et al.*, 2010).

No hay entonces tal contradicción ni tal paradoja en el concepto de homología de función. Si nos preguntamos por qué una determinada especie tiene un determinado rasgo funcional, podemos dar una explicación adaptacionista o analógica de esta función si se identifica su punto de origen evolutivo en el árbol filogenético y se descubren las presiones selectivas que la *diseñaron* en su ambiente ancestral (cf. Brooks y McLennan, 1991, 144). Si ese punto de origen es distinto en otro grupo de comparación que posea un conjunto de rasgos semejantes, entonces en este contexto contrastivo esa semejanza es analógica dado que estos rasgos han evolucionado de manera independiente. Pero podemos responder a esta misma pregunta con una explicación homológica si se mapea el rasgo funcional en la porción del árbol filogenético al que pertenece y se establece que tiene un origen común o filogenia compartida al interior de su propio grupo. La respuesta aquí a la pregunta de por qué estas especies tienen ese rasgo funcional es que lo han heredado de un ancestro común que tenía ese rasgo, es decir, que se trata de una función homóloga.

## 5. EL CASO DE LA EMPATÍA Y EL ALTRUISMO

De Waal (2014) sostiene que ni la compasión ni el altruismo son originalmente actitudes cognitivas en el sentido de que impliquen una previsión de los efectos de la conducta propia en los demás, en la comunidad en general o en uno mismo, o que dependan de procesos superiores, complejos y conscientes. Se trata más bien de procesos simples y, de alguna

9. En el caso de la risa, parece haber, sostienen Davila-Ross *et al.*, 2010, un fuerte vínculo entre forma y función en los grandes simios y en los humanos. Hay, filogenéticamente, dos cambios significativos en el patrón acústico. El primero ocurre en el grupo de primates que se separan de los orangutanes y, el segundo, ocurre en los humanos respecto de los simios africanos. Estas transiciones se deberían presumiblemente a procesos de adaptación (Davila-Ross *et al.*, 2010).

manera, *automáticos*, que forman parte del mecanismo básico del tipo de empatía que subyace a la cognición social y cuyo origen se remonta a la sincronización corporal y la transmisión de estados anímicos o contagio emocional (de Waal, 2014, 148). De Waal describe la empatía como un mecanismo estructurado en tres capas, una básica y emocional en el centro, y otras dos más complejas que se relacionan con ella como las partes de una muñeca rusa (de Waal, 2011; de Waal, 2014). El núcleo de esta estructura es la correspondencia de estados, que de Waal describe como un *mecanismo de percepción acción*, automático e involuntario, que permite a los animales de varias especies tener acceso al estado emocional que perciben en otros individuos de su grupo (de Waal 2007, 64). El proceso de sentir empatía en este caso ocurre, bajo esta propuesta, porque la percepción de conductas y estados del objeto observado (otro individuo del grupo) activa representaciones neurales del sujeto-observador (el que siente empatía). El *contagio emocional* es el correlato conductual de este mecanismo que es posible observar, como se ha dicho, en distintas especies: contagio del miedo, de la risa y del llanto. La *preocupación por los demás* y la *adopción de perspectiva* son las capas que completan el mecanismo básico de la empatía, cada una con su correspondiente manifestación conductual: el *consuelo* y la *ayuda orientada*. Se trata de niveles cognitivamente superiores de empatía, pero que no podrían existir sin la base corporal y emocional de la *correspondencia de estados*, es decir, sin la capacidad de sentirse afectado por lo que ocurre con otros (de Waal, 2007, 66). El consuelo (altruismo socialmente motivado) y la ayuda orientada (altruismo intencional) se observan en unas pocas especies como los grandes simios, los delfines y los elefantes, aunque el consuelo sólo se ha documentado en grandes simios (de Waal, 2007, 60)<sup>10</sup>.

10. De Waal (2007) distingue el altruismo meramente funcional del altruismo motivacional e intencional. La conducta de una abeja que muere al atacar a un extraño es funcionalmente altruista porque aunque supone un coste para ella y un beneficio para la colmena, no está motivada por la colmena (de Waal, 2007, 221). En cambio, el comportamiento de los simios, y otros primates, que se orienta a ayudar a congéneres débiles o heridos, responden a impulsos socialmente motivados (de Waal, 2014, 43; de Waal, 2007, 220-221). El altruismo funcional difiere de los casos cognitivamente superiores de altruismo porque no responde a intenciones conscientes que tengan en cuenta los beneficios para el otro (ayuda orientada) o para sí mismos, como recompensas o retribuciones futuras. El altruismo intencional (ayuda orientada) podría limitarse a humanos, simios y unos pocos animales de grandes cerebros, mientras que el altruismo motivado por expectativas de retribución futura es muy infrecuente fuera de nuestra especie (para una taxonomía del altruismo animal cf. de Waal, 2007, 223).

A este respecto, Philip Kitcher ha sostenido que estos niveles superiores de altruismo psicológico no parecen tener la continuidad evolutiva que requiere el rechazo de de Waal a la teoría de la capa: “la emergencia de la capacidad para la orientación normativa y el auto-control, la habilidad para hablar y discutir potenciales recursos morales con otros”, requieren pasos evolutivos significativos que ponen en duda la tesis de la continuidad (Kitcher, en de Waal 2006, 139-140)<sup>11</sup>. De Waal ha respondido a esta crítica advirtiendo que la mayoría de las veces damos justificaciones *ex post* a nuestros comportamientos altruistas que sin embargo resultan en realidad “automáticos e intuitivos” ajustándose al “patrón común en los humanos de que el efecto precede a la cognición” (de Wall, 2007, 221). Esto es correcto y conduce a una exploración acerca de la naturaleza del carácter moral y su relación con la libertad y la racionalidad. Pero desde la tesis de la estructura jerárquica-desconectada de las homologías de comportamiento es posible también responder a esta crítica. El que exista efectivamente una homología entre, por ejemplo, el altruismo del chimpancé y del humano no exige que el nivel inferior que hace posible biológica y psicológicamente este patrón de comportamiento sea también homólogo. Tal como ocurre en las homologías antes referidas, la evolución pudo haber modificado de manera importante el sustrato que hace posible el patrón de conducta, pero este patrón constituye el mismo rasgo sobre la base de que hay continuidad filogenética, de que ese patrón se originó solo una vez en el ancestro común de ambas especies. Por cierto, podría no haber un único origen. Podría revelarse que el altruismo y la empatía evolucionaron independientemente en los primates no humanos y en los humanos. Pero esta refutación no puede surgir de un diagnóstico filosófico acerca del fundamento de la moralidad humana, sino de una investigación empírica acerca de la evolución de los primates y también de los fundamentos epistémicos de la inferencia filogenética.

## 6. CONCLUSIONES

Hemos mostrado que el uso crítico y heurístico de la antropomorfización defendido por de Waal en el estudio de la conducta social de primates es prescindible y, por esta razón, debería ser reemplazado por una estrategia explicativa en donde la homología funcional ocupe un lugar

11. Entre estos recursos morales o normativos, que “definen la manera en que vivimos”, Kitcher menciona reglas, historias, mitos e imágenes (Kitcher, en de Waal, 2006, 136).

central junto con el criterio de identidad de rasgo biológico que está contenido en el concepto de homología darwiniana. La homología de función es defendible sin abandonar la noción filosófica de la función como el efecto seleccionado y sin desintegrar la homología funcional en términos de la distinción entre actividad y uso con la que Love pretende resolver la supuesta contradicción contenida en el concepto de homología de función. Esta apariencia de contradicción puede disiparse haciendo explícito el carácter contrastivo de la explicación homológica. La crítica de Love se ha revelado entonces como meramente epistémica, su idea de que el significado apropiado de la homología de función es el de actividad y no el de uso o del rol que tiene esta actividad, está motivada por la dificultad de identificar con certeza ese rol como una homología. Pero Love no advierte que la posibilidad de que tales usos sean realizables por distintas funciones de actividad hace asimismo que la homología de actividad carezca de la estabilidad filogenética que él le ha atribuido. Sostenemos que hacer explícita esa estructura jerárquica y la posibilidad de realizabilidad múltiple de patrones de comportamiento social que cumplen funciones homólogas, le permitiría a de Waal responder de manera efectiva a los críticos que niegan su tesis de la continuidad evolutiva de la moral sobre la base de que el substrato biológico y mental humano es suficientemente excepcional para descartar esa continuidad. Si la empatía y el altruismo pueden ser realizables sobre bases biológicas y psicológicas diversas y aun así ser homólogas respecto de conductas semejantes en la familia de primates con la que estamos más estrechamente relacionados, entonces el argumento en contra de la continuidad evolutiva de estas categorías construido sobre la base de la excepcionalidad humana pierde la fortaleza que tradicionalmente ha reclamado.

## REFERENCIAS BIBLIOGRÁFICAS

- Brigandt, Ingo (2007). Typology now: homology and developmental constraints explain evolvability. *Biol Philos*, 22, 709-725. <https://doi.org/10.1007/s10539-007-9089-3>
- Brooks, Daniel y McLennan, Deborah (1991). *Phylogeny, Ecology, and Behavior. A research Program in Comparative Biology*. Chicago: The University of Chicago Press. <https://doi.org/10.2307/1446122>
- Darwin, Charles (1859). *On the Origin of Species*. Londres: John Murray.
- Darwin, Charles (2009[1890]). *The Expression of the Emotions in Man and Animals*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511694110>



- Darwin, Charles (2010[1872]). *El origen de las especies*. Madrid: Edaf.
- Davila-Ross, Marina, Owren, Michael J. y Zimmermann, Elke (2010). The evolution of laughter in great apes and humans. *Comm. Integrat. Biol.*, 3, 191-194. <https://doi.org/10.4161/cib.3.2.10944>
- de Waal, Frans (1988). The Communicative Repertoire of Captive Bonobos (*Pan paniscus*), Compared to That of Chimpanzees. *Behaviour*, 106(3/4), 183-251. <https://doi.org/10.1163/156853988X00269>
- de Waal, Frans (1999). Anthropomorphism and Anthropodenial: Consistency in Our Thinking about Humans and Other Animals. *Philosophical Topics*, 27(1), 255-280. <https://doi.org/10.5840/philtopics199927122>
- de Waal, Frans (2006). *Primates and Philosophers. How Morality Evolved*. New Jersey: Princeton University Press. <https://doi.org/10.1515/9781400830336>
- de Waal, Frans (2007). *Primates y filósofos. La evolución de la moral del simio al hombre*. Barcelona: Paidós.
- de Waal, Frans (2011). *La edad de la empatía. ¿Somos altruistas por naturaleza?* Barcelona: Tusquets.
- de Waal, Frans (2014). *El bonobo y los diez mandamientos. En busca de la ética entre los primates*. Barcelona: Tusquets.
- de Waal, Frans (2021). *¿Tenemos suficiente inteligencia para entender la inteligencia de los animales?* México: Tusquets.
- de Waal, Frans (2022). *La política de los chimpancés*. Madrid: Alianza.
- Dennett, Daniel (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. London: Penguin Books.
- Ereshefsky, Marc (2007). Psychological categories as homologies: lessons from ethology. *Biol Philos*, 22, 659-674. <https://doi.org/10.1007/s10539-007-9091-9>
- Ereshefsky, Marc (2012). Homology thinking. *Biol Philos*, 27, 381-400. <https://doi.org/10.1007/s10539-012-9313-7>
- Fridlund, Alan (1984). *Human Facial Expression. An Evolutionary View*. San Diego: Academic Press.
- Ghiselin, Michael (2005). Homology as a relation of correspondence between parts of individuals. *Theory in Biosciences*, 124(2), 91-103. <https://doi.org/10.1007/BF02814478>
- Ginnobili, Santiago (2022). El origen del rubor. expresión de las emociones, razas y antiesclavismo en la obra darwiniana. *Revista Culturas Científicas*, 3(1), 20-43. <https://doi.org/10.35588/cc.v3i1.5493>
- Hall, Brian K. (1999). *Evolutionary Developmental Biology*. Kluwer Academic Publishers, Dordrecht, Netherlands. <https://doi.org/10.1007/978-94-011-3961-8>

- Hall, Brian K. (Ed.) (1994). *Homology. Hierarchical basis of comparative biology*. San Diego: Academic Press.
- Kant, Immanuel (1991[1800]). *Antropología en sentido pragmático*. Madrid: Alianza.
- Love, Alan C. (2007). Functional homology and homology of function: biological concepts and philosophical consequences. *Biol. Philos.*, 22, 691-708. <https://doi.org/10.1007/s10539-007-9093-7>
- Owen, Richard (1843). *Lectures on the Comparative Anatomy and Physiology of the Invertebrate Animals, Delivered at the Royal College of Surgeons, in 1843*. London: Longman, Brown, Green, and Longmans. <https://doi.org/10.5962/bhl.title.6788>
- Owen, Richard (2012 [1849]). *Discurso sobre la naturaleza de las extremidades*. Oviedo: KRK.
- Papineau, David (2017). *Teleosemantics*. En D. L. Smith (Ed.) *How biology shapes philosophy* (pp. 95-120). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781107295490.006>
- Razeto-Barry, Pablo y Frick, Ramiro (2011). Probabilistic causation and the explanatory role of natural selection. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 42(3). <https://doi.org/10.1016/j.shpsc.2011.03.001>
- Roffé, Ariel J. (2020). Dynamic homology and circularity in cladistic analysis. *Biol Philos*, 35, 21. <https://doi.org/10.1007/s10539-020-9737-4>
- Schmidt, Karen L. y Cohn, Jeffrey F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *Am J Phys Anthropol.*, Suppl 33, 3-24. <https://doi.org/10.1002/ajpa.20001>
- Seyfarth, Robert, Cheney, Dorothy y Marler, Peter (1980). Monkey Responses to Three Different Alarm Calls: Evidence of Predator Classification and Semantic Communication. *Science* 210, 801-3. <https://doi.org/10.1126/science.7433999>
- Torres, Julio (2021). Homología y teleosemántica: el problema de la identidad del rasgo biológico. *Tesis (Lima)*, 14(19), 601-615. <https://doi.org/10.15381/tesis.v14i19.21931>

eISSN: 1989-3612

DOI: <https://doi.org/10.14201/art2024.31271>

# DEL TIMING A LA PLANIFICACIÓN: CONTROL TEMPORAL DE LA CONDUCTA EN ANIMALES NO-HUMANOS

## *From Timing to Planning: Temporal Control of Behavior in Non-Human Animals*

Mauro ZAPATA 

*Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades*

*Secretaría de Ciencias y Tecnologías (SeCyT-UNC), Argentina*

*mzapata@uc.cl*

Recibido: 31/03/2023

Revisado: 15/08/2023

Aceptado: 08/11/2023

**RESUMEN:** La cognición orientada a futuro, el conjunto de procesos cognitivos que incorporan implícita o explícitamente consideraciones de estados futuros, es un tópico que ha adquirido creciente interés en los estudios de cognición animal. Mediante el análisis de literatura sobre aprendizaje asociativo y estudios en etología, justificaremos la atribución de cognición prospectiva a distintas especies, distinguiendo dos formas no excluyentes de cognición orientada a futuro, por una parte, la planificación, y por otra el timing: la capacidad cognitiva de adaptar el comportamiento frente a las regularidades temporales del entorno. Aplicaremos esta distinción al trabajo llevado a cabo por Frans de Waal en chimpancés.

*Palabras clave:* cognición orientada a futuro, cognición animal, condicionamiento, demora, uso de herramienta.

**ABSTRACT:** Future oriented cognition, the group of cognitive processes that incorporate explicit or implicit considerations of future states, is a topic that has received growing interest within animal cognition studies. Through the analysis of associative learning literature and ethological studies, we will justify the attribution of prospective cognition to different species, by way of distinguishing two non-exclusive forms of future oriented cognition: on the one hand planning, on the other hand timing: the cognitive capacity to adjust behavior to temporal regularities in the environment. We will apply this distinction to the work of Frans de Waal on chimpanzees.

*Keywords:* future-oriented cognition, animal cognition, conditioning, delay, tool use.

## 1. INTRODUCCIÓN

En formas diversas, unas más sofisticadas que otras, animales humanos y no-humanos cotidianamente desarrollan actividades para su supervivencia. Seleccionar alimentos para el consumo (Phillips y Lancelotti, 2014), integrarse a grupos de pares (McClung *et al.*, 2020) y usar y construir herramientas (C. M. Sanz *et al.*, 2013) son solo algunas actividades complejas que compartimos con nuestros parientes evolutivos cercanos. Todas formas de interactuar con el entorno que requieren algún grado de coordinación temporal o anticipación.

Ante una acción o conducta desplegada por un animal, el componente temporal del entorno en que se desenvuelve juega un rol específico en el desempeño exitoso de la tarea. Así, se volverá fundamental para su supervivencia la capacidad para adaptar el comportamiento frente a las regularidades temporales del entorno. Dicha capacidad es conocida como *timing* en la literatura asociada (Vasconcelos *et al.*, 2017), especialmente si se trata de intervalos arbitrarios o artificiales generados en contextos de laboratorio<sup>1</sup>.

Los ajustes que un organismo realiza a partir de ciertas regularidades pueden variar en una escala microbiológica y de corta duración, por ejemplo, en la detección de los ciclos de noche y día (Panda *et al.*, 2002), hasta comportamiento observable a gran escala y de larga duración, cuando se trata de ajuste a los cambios en las estaciones del año (Gwinner, 1986). Se puede sostener la importancia de contar con un mecanismo sensible

1. En estos contextos, la voz "*timing*" se utiliza de forma específica como abreviatura de la etiqueta "*interval timing*".

a las regularidades temporales a partir de la presión selectiva que ejercen los ciclos naturales sobre las criaturas que los sufren.

En el presente ensayo buscaremos justificar la atribución de distintas formas de cognición orientada al futuro a animales, entendida como el conjunto de ciertos procesos cognitivos que incorporan implícita o explícitamente consideraciones de estados futuros (Bubić y Abraham, 2014). Evaluaremos evidencia empírica de *timing* en animales en entornos controlados de laboratorio, distinguiéndolos de conductas más sofisticadas que evidencian planificación reportadas en medio libre. Por último, analizaremos a partir de la distinción propuesta entre *timing* y planificación, la evidencia aportada por el primatólogo Frans de Waal al estudiar una colonia de chimpancés en el zoológico de Arnhem. Así, podremos apreciar cómo la anticipación de un estado futuro opera en cada uno de estos casos, habilitando la atribución de cognición orientada a futuro.

## 2. *TIMING*, INHIBICIÓN DE LA DEMORA E INTERVALOS FIJOS

Aunque altamente controvertido (Allen-Hermanson, 2005; Starzak, 2017), el llamado Canon de Morgan sigue siendo una regla metodológica ampliamente adoptada a la hora de investigar, en etología y psicología comparada, capacidades de las cuales se sospecha que puedan ser cognitivas. Así, el Canon (Morgan, 1894) que llama a no interpretar en términos de habilidades psicológicas superiores comportamientos que pueden ser explicados a partir de habilidades inferiores (en términos evolutivos o de desarrollo), invita a abstenernos de postular representaciones temporales, de momento, para partir *desde abajo* y escalar hacia una explicación cognitiva.

Como advirtió en su momento Allen (Allen, 2018), los excesos cometidos en el análisis o tratamiento anecdótico del comportamiento animal, atribuidos famosamente a Romanes (Romanes, 1888) llevaron a preferir las explicaciones por aprendizaje (asociativo)<sup>2</sup> como la alternativa deflacionada a las explicaciones cognitivistas, respetando así el Canon.

Pero será justamente en el aprendizaje asociativo que encontraremos rasgos iniciales de procesos anticipatorios: lo que llamamos *timing* en un sentido técnico, esto es, un ajuste en la respuesta del animal a partir de la

2. Aquí aprendizaje es simplemente definido como la modificación del comportamiento como función de la experiencia.

consideración de las propiedades o relaciones temporales de los eventos ocurridos durante el condicionamiento. Podremos encontrar estos elementos tanto en el condicionamiento clásico como operante.

Recordemos que en el *condicionamiento clásico* (Pavlov, 1927), tratamos con un programa que intenta obtener una respuesta conductual a partir de un estímulo inicialmente neutro. En el ejemplo tradicional se busca obtener la respuesta de salivación de un animal (respuesta incondicionada, en adelante RI) que normalmente ocurre en presencia de comida (estímulo incondicionado, EI), pero esta vez a partir de un estímulo distinto, neutro, como puede ser una campana, otro sonido específico o una luz (estímulo condicionado, EC). La salivación resultante en presencia del EC recibe el nombre de respuesta condicionada (RC), y fue utilizada como medidor de que el vínculo por asociación entre ambos estímulos (EI y EC) dio resultado.

Dentro de este paradigma, resulta de especial interés un subtipo de condicionamiento, el *condicionamiento demorado*. En este tipo de condicionamiento, el EC es presentado antes que el EI y el término del EC coincide con la ocurrencia del EI o se da en el transcurso de éste (Froufe Torres, 2011). Esta configuración permitió inicialmente detectar un fenómeno al realizar ciertos ajustes en la duración del intervalo entre la presentación del EC y la del EI. Al aumentar la duración del tiempo en que está presente el EC sin que se haga presente aún el EI, se produce el siguiente resultado: al inicio del entrenamiento la latencia se acorta (i.e. el tiempo que transcurre entre el inicio del EC y la exhibición de la RC), pero a medida que continúa el entrenamiento, esta se alarga hasta el momento en que ocurre el EI.

Este hallazgo fue denominado en la literatura como *inhibición de la demora* (Pavlov, 1927), y el término ha sido utilizado para describir escenarios de condicionamiento demorado en los que hay una presencia mínima o nula de RC al inicio del período largo de un EC, como también para situaciones en las que la RC presente al inicio de un período largo de EC disminuye en la medida que el entrenamiento avanza. Esto último, a su vez, ha sido interpretado como sugiriendo que el tramo inicial del EC adquiere un control conductual que decrece en la medida que avanza el aprendizaje de la relación temporal entre ambos estímulos (EC y EI) (Escobar *et al.*, 2015).

Así, la literatura asociada ha entendido que podemos evaluar si nos encontramos frente a un caso de inhibición de la demora atendiendo a las siguientes características (Rescorla, 1967):

- a. Una disminución (con el condicionamiento) en la respuesta evocada por el inicio del EC.
- b. Confinamiento de la RC hacia la parte final del EC.
- c. Supresión de la tasa de respuesta en la parte inicial del EC.

Dicho de otro modo, en la inhibición de la demora el animal aprende que el EI no ocurrirá sino hasta que un buen tiempo haya transcurrido desde que inició el EC y en consecuencia inhibe la RC en el tramo inicial (Froufe Torres, 2011), guardándola para el tiempo en que debiera presentarse el EI según se mostró en intentos anteriores.

Habría entonces una explicación de la inhibición de la demora en la que no solo se establece un vínculo asociativo entre estímulos (EC y EI) y estímulos con respuestas (EC y RC), sino que nos encontramos en un escenario en el que el animal almacena la duración de los intervalos entre un evento y otro y recupera esta información para el proceso de toma de decisión que determina la inhibición (o no) de la conducta observable (RC). El animal aprende, entre otros, las relaciones temporales entre los eventos que ocurren dentro del entrenamiento por asociación (*intervalos temporales*), aprendiendo a la vez su propiedad recíproca, esto es, información acerca de la *tasa de ocurrencia de un evento*, siendo el aprendizaje de ambas señalado como el proceso central en los casos de condicionamiento clásico e instrumental (Gallistel y Gibbon, 2000).

Si bien la inhibición de la demora aparece como un ejemplo relevante de *timing*, los desarrollos experimentales siguieron más bien el paradigma experimental del condicionamiento operante diseñado por Skinner (Skinner, 1938). En este paradigma, la noción clave será la de respuesta operante.

El interés de Skinner era medir en contextos controlados de laboratorio conductas que fueran representativas del comportamiento natural de las criaturas en medio libre. A este efecto, Skinner diseña un dispositivo, un espacio cerrado y reducido en el que el animal puede circular libremente dentro de sus límites, debiendo activar un dispositivo (normalmente una palanca en el caso de las ratas) para obtener una recompensa, normalmente alimento (refuerzo), no estando limitado a un número de veces determinado (tasa de respuesta) como tampoco a momentos específicos en que la conducta debe ser llevada a cabo. La respuesta operante entonces es la conducta útil en el contexto del experimento (en el caso de la rata presionar la palanca), respuesta definida en términos del efecto que el comportamiento tiene en el ambiente.

La manipulación de la palanca (en el caso de la rata) recibe el nombre de respuesta operante, aunque no sea inicialmente una respuesta a ningún estímulo (al menos en el mismo sentido en que la RC si es una respuesta al EC para el condicionamiento clásico). Así, en el paradigma Skinneriano “*respuesta*” se vuelve sinónimo de “unidad medible de comportamiento” (Wearden, 2016).

Guiado en su diseño por la ley del efecto (Thorndike, 1905), la medida del condicionamiento operante es la tasa de respuesta<sup>3</sup> y el patrón de respuesta<sup>4</sup>. Estos elementos son analizados en el contexto de la duración completa del proceso de entrenamiento bajo el diseño de la llamada *agenda de refuerzo*, i.e. la relación entre el tiempo transcurrido, la conducta desplegada por el animal y la frecuencia en la entrega del refuerzo o recompensa.

Nos centraremos en una clase especial de agenda de refuerzo conocida como agenda de *intervalo fijo* (IF). En este diseño, solo la primera respuesta es reforzada y esto da inicio a un reloj que cuenta por  $t$  segundos. Al finalizar el tiempo, la próxima respuesta es reforzada y el reloj se reinicia. Luego de un tiempo de entrenamiento se logra apreciar que los animales exhiben control temporal de sus respuestas, siendo esta una función del tiempo transcurrido (intervalo) (Whitaker *et al.*, 2003).

Lo anterior se evidencia por medio del patrón de respuesta, y es aquí donde encontramos particularidades. La entrega de un refuerzo provoca una pausa-post-refuerzo (PPR) (Lowe *et al.*, 1979). Una vez concluida la PPR, el animal exhibe nuevas respuestas. Ahora bien, estas respuestas luego de la PPR instancian uno de dos patrones.

En primer lugar, pueden exhibir un patrón en el que se retoma la respuesta de manera paulatina, acelerando gradualmente, llegando a su máximo justo antes del siguiente refuerzo. A este patrón se lo denomina “*escalopado*” (Dews, 1978).

El segundo patrón de respuestas que puede ocurrir luego de la PPR es el llamado “*patrón de dos estados*”. En este patrón distinguimos un primer estado en que la tasa de respuesta es baja y constante luego del refuerzo, de un segundo estado que ocurre en un punto variable luego del refuerzo. En este punto hay un paso abrupto a una tasa alta constante. El punto de transición ocurre alrededor de dos-tercios del camino

3. Número de respuestas por unidad de medida.

4. La distribución específica que tienen las respuestas en el intervalo de tiempo medido.



del intervalo (que va de refuerzo a refuerzo) y en términos técnicos se puede señalar que la tasa de respuesta en el segundo estado es una función creciente y negativamente acelerada respecto de la tasa de refuerzo en el segundo estado (Schneider, 1969). Los dos estados son entendidos como estados discretos, en que el primero presenta nula o casi nula respuesta, mientras que en el segundo consta el aumento abrupto a una tasa alta y constante.

Si consideramos ambos casos, las similitudes en las propiedades de los patrones de las agendas de intervalo fijo y la inhibición de la demora son patentes. Si recordamos las exigencias con que Rescorla caracterizó la inhibición de la demora, tenemos que los puntos (b) y (c) se mantienen para uno y otro caso haciendo los ajustes correspondientes. Reformulando estas exigencias para el caso de FI tenemos las siguientes proposiciones:

(b!) Confinamiento de la respuesta hacia la parte final del intervalo (o el segundo estado en el patrón de dos estados).

(c!) Supresión de la tasa de respuesta en la parte inicial del intervalo (o en el primer estado en el patrón de dos estados).

Estimamos que esta equivalencia en las propiedades temporales de los patrones de respuesta en ambos casos habilita la atribución de timing o control temporal de las respuestas a los sujetos en cuestión en cada uno de los estudios.

Sobre este punto cabe hacer presente que los estudios que dieron cuenta del patrón de dos estados en IF fueron realizados con palomas blancas Carneau (Schneider, 1969), pero Sherman obtuvo resultados similares en ratas (Sherman, 1959). Por su parte, los experimentos que reivindicaban el patrón escalopado como una ocurrencia usual en experimentos de IF fueron realizados con macacos Rhesus (Dews, 1978), pero evidencia posterior arrojó resultados similares en diversas especies, incluidas gatos, ratas, palomas, tórtolas y tortugas marinas (Lejeune y Wearden, 1991). Por último, los experimentos de inhibición de la demora suelen llevarse a cabo sobre perros (Rescorla, 1967) y ratas (Zieliński *et al.*, 1995) principalmente.

Hasta aquí hemos mostrado cómo los casos de *timing* mediante los fenómenos de inhibición de la demora y la agenda de intervalo fijo reflejan una capacidad para anticipar el momento de entrega del refuerzo y ajustar su conducta.

### 3. PLANIFICACIÓN

Para comprender el lugar que ocupa la planificación dentro de la vida mental de chimpancés y otras especies, convendrá tener presente que recientemente ha sido propuesta la tesis del cerebro o mente como una máquina de hacer predicciones (Bubic *et al.*, 2010; Dennett, 1992; Gärdenfors, 1995). En un sentido amplio, en el procesamiento predictivo que realiza el cerebro se utiliza y genera información no solo del pasado y presente, sino de estados futuros, tanto del cuerpo como del entorno (Bubic *et al.*, 2010). Esta idea se encuentra a la base de lo que hemos denominado aquí cognición orientada a futuro.

Distinguiremos esta caracterización general del procesamiento predictivo en sentido amplio, de la hipótesis específica del procesamiento predictivo (o *predictive coding*) según la cual distintas áreas del neocórtex implementan un modelo generativo del mundo que es aprendido desde los inputs sensoriales. Luego las conexiones de *feedback* con dirección *top-down* transmiten predicciones de actividad de nivel bajo, mientras que las conexiones con dirección *bottom-up* transmiten los errores entre las predicciones *top-down* y la actividad que efectivamente tuvo lugar (Rao y Jiang, 2022).

Una idea similar al procesamiento predictivo es expresada en la llamada "hipótesis de la simulación". Según esta concepción, el pensamiento consiste en una serie de interacciones simuladas con el entorno (Hesslow, 2002). La idea en juego es que al menos algunas formas importantes de pensamiento involucran los mismos procesos que el interactuar con el entorno exterior, con la gran diferencia de que las acciones se encuentran encubiertas (no se externalizan). Estos procesos generarían los mismos inputs sensoriales que aquellos que desencadenan acciones posteriores. Adicionalmente, dichos inputs serían generados mediante mecanismos asociativos (Hesslow, 2012).

La teoría distingue de esta manera tres clases de simulación (Hesslow, 2002, 2012):

1. Simulación del comportamiento, en la que las áreas pre-motoras en el lóbulo frontal son activadas de manera similar a como se activan normalmente para el movimiento, pero la actividad neuronal es interrumpida antes de que tenga ocasión el comportamiento externo.
2. Simulación de la percepción, en la que la corteza sensorial presenta actividad que es provocada por otras áreas, actividad similar a la que tiene lugar cuando hay percepción de estímulos externos.

3. Anticipación, operando mediante mecanismos asociativos, se habilita que actividad perceptual y conductual produzca actividad en las áreas sensoriales del cerebro. Así, una acción simulada puede generar actividad perceptual similar a la que habría ocurrido si la acción hubiese tenido lugar. A su vez, una percepción simulada hace las veces de estímulo para una nueva acción encubierta.

Hesslow es enfático a la hora de señalar que su teoría se encuentra basada únicamente en mecanismos asociativos, por lo que en su estructura no encontraremos símbolos, representaciones o modelos internos (como sí encontramos en la teoría del procesamiento predictivo). Esto, sin embargo, no vuelve a la teoría inconsistente con una que incluya representaciones o modelos, pues estos últimos pueden tener lugar a partir de los vínculos asociativos según se defina “representación” y “modelo” respectivamente<sup>5</sup>.

A pesar de su inspiración asociacionista, la teoría de la simulación nos entrega una imagen clara de cómo podría el cerebro impulsar una cadena de procesos cognitivos que involucren estímulos y acciones que, en la práctica, no están teniendo lugar, pero que podrían llegar a tenerlo (i.e. estados futuros).

Volvamos ahora sobre la cognición orientada a futuro en animales no-humanos. Esta, no solo tiene lugar en contextos de laboratorio como aquellos que corresponden a los condicionamientos clásico y operante. También puede ser observada en contextos ecológicos en los que preocupaciones acerca de la interferencia que pudiera ocasionar el entrenamiento en la modificación de la conducta pueden ser aliviadas. En el medio libre, encontraremos ejemplos de conductas que son explicadas desde la cognición orientada al futuro mediante procesos cognitivos tales como:

1. simulación, tradicionalmente entendida como la construcción de una representación mental de eventos futuros hipotéticos.
2. predicción, esto es, la estimación de la probabilidad de ocurrencia de un evento futuro.
3. intención, el acto mental de establecer una meta u objetivo, y
4. planificación: identificación y organización de pasos para arribar a un estado, evento o resultado futuro (Szpunar *et al.*, 2014).

5. Para una noción relevante de “modelo” ver Holland y Goodman (2003).

Nos centraremos en este apartado en los procesos de planificación, por ser una clase de procesos compleja, que en ocasiones presupone que los demás procesos (simulación, predicción e intención) tengan lugar, pero, adicionalmente, por encontrarse importantemente documentado en primates (C. M. Sanz y Morgan, 2009) y aves (Kabadayi y Osvath, 2017).

En cuanto a las conductas en medio libre que puedan evidenciar planificación, se ha postulado que la búsqueda y recolección de alimento actuó como presión de selección para el desarrollo de habilidades cognitivas entre las que se encuentran habilidades de engaño, uso de herramientas y aprendizaje observacional (McLean, 2001). Es en el despliegue de estas habilidades que encontraremos indicios de planificación.

Para el caso del uso de herramientas, diversos estudios han reportado la conducta de elaboración y transporte de herramientas en chimpancés (Boesch y Boesch, 1984; Goodall, 1964). En el trabajo llevado a cabo por Boesch y Boesch, se estudió el comportamiento de chimpancés del Parque Nacional Tai (Costa de Marfil). La conducta en cuestión fue la elaboración y transporte de martillos y piedras para lograr romper nueces que se encontraban en diversos sitios. La selección de la herramienta ocurría en ausencia del alimento y esta era transportada hasta el sitio de uso, donde sólo fue transportado aquel martillo lo suficientemente fuerte como para romper el tipo de nuez en cuestión (de entre dos tipos, uno blando y uno más resistente). Encontramos planificación no solo en la selección de la herramienta, sino en la elección del sitio de recuperación, tanto de la herramienta como del alimento, tendientes a maximizar la eficiencia de los esfuerzos, no cargando por tramos largos las herramientas más pesadas.

En un segundo caso, los chimpancés del bosque ubicado en el Triángulo Goulougo, en el Congo, discriminaban y planificaban el traslado de dos tipos de herramientas distintas. Unas con punta para poder perforar nidos de termitas sobre la superficie, y otro con forma de caña de pescar para recuperar alimento de sitios poco accesibles. En todos los casos (45 intentos registrados), los sujetos trasladaron únicamente el tipo adecuado de herramienta para la clase de nido que buscaban explotar (C. Sanz *et al.*, 2004). Esto puede ser interpretado como indicio de que los chimpancés están, en términos del proceso cognitivo subyacente, simulando mentalmente una acción futura, en concreto, la de obtener o rescatar el alimento en un lugar distinto a aquel en que se encuentra el sujeto al momento de la simulación. Por otra parte, la representación de los pasos para seleccionar, recuperar y trasladar las herramientas al sitio de uso, constituirían un caso de planificación como definido anteriormente.

Ahora bien, estar frente a un caso de planificación no implica estar frente a un caso de control temporal. Para que la planificación pueda considerarse análoga a los casos de timing estudiados en la sección anterior, se requerirá no solo la evidente orientación a futuro de la conducta, sino también que la modificación o despliegue de la conducta (en este caso de parte del chimpancé) sea sensible a los factores temporales de la tarea en cuestión. Hasta aquí, los factores que especulativamente pueden establecerse como determinantes de la conducta de los sujetos, son consideraciones sobre la función que puede desempeñar una u otra herramienta, la clase del alimento y distancia entre los sitios. Se requerirá información adicional rescatada de manera sistemática acerca de, por ejemplo, la disponibilidad de alimento en determinadas épocas del año, para poder sostener que hay elementos temporales en la configuración de la tarea que están influyendo en la planificación, tornándola así en un caso de control temporal.

Esto ocurre en el caso de un estudio llevado a cabo por Janmaat et al. (2014), quienes parten de la intuición capturada por la llamada hipótesis de la inteligencia ecológica (Milton, 1981), según la cual la diversidad en los alimentos disponibles para una especie y su distribución espacial y temporal constituyeron fuerzas selectivas relevantes para el desarrollo de la complejidad cerebral. Partiendo de la base de que la distribución del alimento puede decir algo acerca de las capacidades cognitivas de quienes los consumen, procedieron a realizar un estudio durante 275 días en una selva del oeste africano. El registro consistió en tomar nota de dónde y cuándo los chimpancés en estudio pasaban la noche y en que lugar luego obtenían comida al día siguiente. La hipótesis para constatar era que, existiendo una diferencia en el tipo de alimento disponible en los diversos sitios, debiera haber una diferencia en el comportamiento de los chimpancés para los sitios donde el alimento parece más rápido y, por tanto, son considerados como sitios más codiciados. Esto, en contraste con los sitios en que abunda alimento que no parece fácilmente. Así mismo, cobró relevancia la distancia a la que se encontraba el sitio de alimento respecto del sitio para pernoctar. Los resultados arrojaron que los chimpancés usualmente partían más temprano (muchas veces antes del amanecer) cuando desayunaban en sitios con comida efímera, sobre todo si se encontraban más lejos. A través de los diversos controles establecidos en el estudio, pudieron concluir que los chimpancés planificaban de forma flexible el horario, tipo y ubicación de sus desayunos en función del declive en la disponibilidad del alimento por temporada.

A diferencia de los estudios sobre traslado y uso de herramientas, vemos que son justamente factores temporales como el tiempo en el que perece un alimento, las horas del día y los tiempos de traslado, los que explican la modificación del comportamiento de estos chimpancés a la hora de despertar y dirigirse al sitio de desayuno. Podemos entonces sugerir que para este caso estamos en presencia de control temporal, evidenciado en la sensibilidad a factores temporales en la planificación del desayuno.

#### 4. EL CASO INTERMEDIO

Hasta aquí hemos revisado conductas relativas al control temporal en entornos de laboratorio como también de planificación en medio libre. Dicho esto, uno de los casos mejor documentados de conductas sofisticadas en homínidos ocupa una posición intermedia. Hablamos del estudio que llevó a cabo Frans de Waal entre los años 1975 y 1981 sobre la colonia de chimpancés del zoológico de Burgers en Arnhem (Países Bajos) (de Waal, 2022). Denominamos este un caso intermedio por las características particulares del espacio en que se encuentra la colonia al interior del zoo. A pesar de ser un zoológico, el espacio en que conviven unos con otros dista mucho de ser un lugar de encierro y exhibición de animales al público. Es un espacio al aire libre, de alrededor de dos acres de extensión, con árboles vivos y troncos secos para trepar, todo rodeado por una fosa con agua. Durante los años reportados por de Waal la colonia albergó un grupo de hasta veinte individuos. Los humanos visitantes quedan apostados a veinte metros del lugar donde se encuentra la colonia, lo que evita conductas disruptivas como alimentar o confrontar a los animales. Esta configuración permite respetar las dinámicas de la colonia como si ocurrieran en medio libre, obteniendo una observación fidedigna de sus interacciones a pesar de encontrarse de cierto modo reclusos. Adicionalmente, permite realizar intervenciones por parte de los cuidadores humanos para resguardar la integridad de los miembros de la colonia por medio de la separación de individuos especialmente violentos a raíz de episodios específicos.

El estudio sobre la colonia de Arnhem fue realizado por un grupo de investigadores que registraron en la segunda mitad de la década del '70 las diversas interacciones sociales entre los integrantes de la colonia. El espectro de interacciones abarca tanto conductas amistosas como hostiles. Dentro de las interacciones amistosas entre chimpancés encontramos la limpieza de unos a otros o "espulgamiento", como también "juego,

coaliciones y saludos” (de Waal, 2022, 194). Entre las conductas hostiles encontramos gritos, manotazos, atropellos, mordeduras y ataques con objetos contundentes (ramas y piedras).

En su libro “La política de los chimpancés”, originalmente publicado en 1982, de Waal entrega los resultados de las observaciones realizadas durante la década anterior, poniendo énfasis en relatar los cambios en la estructura social de la colonia, en términos de los grupos, alianzas formadas y disueltas, así como los individuos que terminan tomando el poder constituyéndose en macho alfa hasta que un nuevo proceso de toma de poder ocurra.

Ya desde un comienzo, de Waal nos entrega distintos ejemplos que permiten reconocer ciertas habilidades cognitivas orientadas al futuro. El autor recuerda el relato entregado por los funcionarios del zoo sobre un tiempo anterior a la llegada de de Waal, quienes presenciaron la fuga de los chimpancés desde su espacio dedicado, hacia el resto del zoo. La hazaña fue realizada colocando ramas contra el muro para ser trepadas, ayudándose unos a otros a escalarla. Este relato fue confirmado personalmente por de Waal quien entrevistó las personas encargadas de la cocina en aquella época, quienes reportaron que a lo menos un chimpancé (“Mamá”, la chimpancé hembra de mayor edad de la colonia) ingresó al casino en busca de comida, encontrando solo una botella de leche chocolatada, que acabó ella sola tranquilamente. La manera en que se efectuó la fuga, colocando ramas de alrededor de cinco metros para superar un muro que asciende a cuatro metros, impresiona “como si el plan hubiese sido acordado de antemano” (de Waal, 2022, 43).

Una muestra adicional de cognición orientada a futuro la encontramos en el caso de Dandy, el único de cuatro machos de la colonia que no participó realmente de las pugnas de poder. Siendo el más joven, debió desarrollar su propia manera de lograr sus objetivos, en concreto, de satisfacer sus necesidades más básicas de que otros chimpancés se le acerquen y lo espulguen. Así, ideó una forma de comunicación gestual para este efecto, tomando la parte superior de su brazo izquierdo con la mano derecha como señal, cada vez que pretende que otro chimpancé se acerque a limpiarlo (de Waal, 2022, 55). Apreciamos en este caso un ejemplo de intención, como forma de cognición orientada a futuro.

Una explicación alternativa que prescinda de las intenciones descansaría o bien en un mecanismo rígido propio de la especie o bien en un proceso de aprendizaje asociativo. Si fuera lo primero, la conducta de Dandy debiera entenderse como una reacción a un conjunto específico de circunstancias que concurren en un momento dado, y no podría ocurrir que estas concurren y sin embargo no tenga lugar la conducta

(gestos). La especificación de estas circunstancias tendría que ser lo suficientemente fina como para dar sentido al hecho de que Dandy no se encuentre realizando estos gestos constantemente, y que los gestos cesen una vez que da inicio la conducta de espulgamiento. Si bien posible de ser entregada, esta explicación del comportamiento gestual de Dandy ya no aparece como una alternativa más simple. Adicionalmente, esperaríamos encontrar comportamiento idéntico en más de un individuo de la especie y esto contrasta con la evidencia. Para un aprendizaje por asociación, se requeriría que Dandy haya sido expuesto con anterioridad al par de eventos (gesto-específico, espulgamiento) a fin de que pueda percibir e internalizar la relación contingente entre ambos. Esto también contrasta con la evidencia, ya que no se constató este gesto en ningún otro individuo de la colonia. El gesto paradigmático con el que el resto de la colonia inicia una interacción de espulgamiento es simplemente estirar el brazo con la palma abierta (*"ofrecer la mano"*) (de Waal, 2022, 55). Ahora bien, el gesto de Dandy presenta una especificidad mayor que el gesto de *"ofrecer la mano"*, ya que este último es observado seguido de conductas tan diversas como: recibir comida, mantener contacto corporal y recibir apoyo durante un enfrentamiento (de Waal, 2022, 56); mientras que el de Dandy ocurre únicamente seguido del espulgamiento. Lo anterior torna más plausible la explicación en la que la conducta gestual de Dandy es una novedad conductual realizada con la finalidad en mente de ser espulgado.

Un tercer ejemplo entregado por de Waal, son las conductas de confiscación. Para entender el carácter orientado a futuro se debe tener presente la manera en que se producen las interacciones hostiles entre chimpancés. Normalmente, realizan exhibiciones de fuerza o poderío, emitiendo ruidos, dando saltos y poniendo sus pelos en punta, logrando que parezca que su volumen corporal es mucho más grande que el que realmente tienen (de Waal, 2022, 58). Estas conductas se encuentran dirigidas comúnmente a un rival y, en ocasiones, el chimpancé que realiza la demostración de fuerza traerá consigo un arma, esto es, piedras en la palma de la mano o palos. Con esto no pretenden amedrentar al rival sino concretar una agresión. Es en este escenario que interviene comúnmente una hembra quien abre lentamente los dedos de la mano donde cargan el arma, y la retira de sus manos. Si tenemos presente que la hembra luego de confiscar el objeto contundente no realiza ninguna acción adicional con el mismo, resulta plausible interpretar esta conducta como motivada por el propósito de evitar la agresión futura. Para que esto ocurra, la hembra que no intervenía en el conflicto hasta percatarse de la presencia de las piedras debe simular el evento futuro, esto



es, generar una representación del evento consistente en la agresión que un chimpancé realiza contra otro arrojando o golpeando con las piedras.

Pongamos ahora la atención en una de las dinámicas centrales tratadas en el estudio, esto es, las dinámicas de toma de poder. Durante el período reportado por de Waal, tres machos disputaron el poder entre ellos. Yeroen, el macho alfa original, se convierte en líder de la colonia a su llegada. Luit, quien llega con Yeroen cuando son incorporados a la colonia, disputa el poder justamente de las manos de Yeroen en la primera lucha por el poder. Por último, está Nikkie, el tercer macho cuyo apoyo resultó crucial para decidir las tomas de poder en uno u otro sentido.

Los enfrentamientos ocurren cuando una de las partes (en este caso Luit) inicia demostraciones de fuerza provocando al rival. Si el rival responde ante la amenaza, comienza una persecución y enfrentamiento físico entre las partes. Lo interesante es que, mientras que la batalla física se decide por la fuerza y agilidad de los chimpancés, la pelea política se decide por el respaldo que tenga una de las partes del resto de la colonia (las hembras). Así, el estudio pudo comprobar que cuando se aproximaba un enfrentamiento, Yeroen aumentó las conductas amistosas de juego y espulgamiento que normalmente ocupaban sólo un treinta por ciento de sus interacciones a prácticamente el doble. De Waal interpreta esta conducta como un intento de ganar el favor de las demás integrantes de la colonia. Esta interpretación se ve reforzada por la conducta de respuesta llevada a cabo por Luit, quien al constatar estas interacciones se dirige rápidamente a interrumpirlas, haciendo demostraciones o derechamente atacando. Esta conducta recibe el nombre de “intervención de separación” (de Waal, 2022, 158).

Denominarlas de esta forma sugiere que de una u otra manera, los involucrados comprenden el valor que esas conductas tienen a la hora de asegurar respaldos de la comunidad para los enfrentamientos. Por otra parte, tanto la conducta para ganar el favor como la conducta de separación ocurren con una antelación importante (de días) al enfrentamiento, lo que resulta indiciario de una capacidad anticipatoria. Así descritos, resulta difícil hacer sentido de la conducta de hembras y machos sin que ellos tuviesen un mecanismo cognitivo para estimar o anticipar el resultado de una acción determinada, como puede ser el confiscar un arma, generar conductas amistosas o expresar conductas hostiles, previendo la posibilidad de un enfrentamiento, ya sea para evitar el mismo, o para facilitar su ocurrencia.

Una conducta orientada a futuro adicional se da justamente en el período que sigue inmediatamente a un enfrentamiento. Nikkie, en un enfrentamiento con Yeroen donde el primero resultó vencedor, vio sus

números reducidos por culpa de los esfuerzos de Yeroen por maximizar sus interacciones amistosas con las hembras del grupo en el tiempo anterior al enfrentamiento. Una vez concluida la batalla, Nikkie como vencedor espera media hora, tras lo cual realiza conductas amistosas con Yeroen. De Waal interpreta que los chimpancés de la colonia pueden leer de una u otra forma la tensión en el ambiente y sería esta la razón por la que se apresuran en dar cierto nivel de estabilidad al grupo mediante conductas amistosas a modo de reconciliación. Adicionalmente, Nikkie hace todo lo contrario con las hembras que se levantaron en su contra y opta por castigarlas mediante agresiones. Esta conducta es interpretada como desincentivando futuros levantamientos de las hembras o alianzas de ellas con sus contrincantes (de Waal, 2022, 237). Nuevamente, nos encontramos frente a la anticipación de un evento futuro y el rauda actuar del macho alfa para evitarlo.

Ahora bien, el mismo autor reconoce que a la base de esta diversidad conductual se encuentran algunos mecanismos más o menos sofisticados que no poseen necesariamente un carácter temporal. El autor señala como relevante en primer lugar la capacidad de reconocerse individualmente los unos a los otros. Esto evita la necesidad de tener que reafirmar las posiciones de dominación en cada interacción o reencuentro con otro miembro del grupo o colonia. Sin esta capacidad no se podría otorgar un mínimo de estabilidad a las estructuras sociales. Pero se considera aún más importante la llamada conciencia triádica, esto es, la “capacidad de percibir las relaciones sociales que se dan entre otros individuos y formar relaciones triangulares variadas” (de Waal, 2022, 292). Esta capacidad fue la que permitió que se ejerciera un control tripartito por parte de Yeroen, Luit y Nikkie en la colonia en Arnhem, habilitando el monitoreo de las relaciones de los dos restantes entre ellos y con otros ejemplares de la colonia. El monitoreo a su vez dio lugar a las ya mencionadas intervenciones de separación, y la conciencia triádica aparece como la capacidad responsable de la posibilidad de coaliciones, rencillas y reconciliaciones.

Habiendo verificado la riqueza conductual del estudio realizado por de Waal, llama la atención que al explorar el índice analítico en “La política de los chimpancés” en busca de la palabra planificación, esta aparezca solo dos veces. Ciertamente esto es reflejo del foco de la investigación en las estructuras sociales y sus dinámicas de poder y nada resta a la aproximación que hemos dado a la misma evidencia en este trabajo. Si algo, es testimonio de la influencia que ha tenido la reflexión sobre la cognición prospectiva en las décadas recientes y cómo este enfoque arroja nueva luz sobre información reunida desde que iniciaron los diversos programas

de investigación en primates en medios de laboratorio, medio libre, santuarios y zoológicos (McEwen *et al.*, 2022).

Pero si queremos ser justos con el autor, lo cierto es que, de forma visionaria se ocupa de la cuestión que nos convoca, esto es, la cuestión acerca del pensamiento sobre el futuro. Es en el acto de preguntarse sobre una eventual racionalidad en las decisiones de alianza entre los machos de la colonia que de Waal evalúa la orientación a futuro de sus conductas. Entendiendo que “una elección racional está basada en la estimación de las consecuencias” (de Waal, 2022, 303), la pregunta es si acaso Yeroen pensaba en el futuro a la hora de decidir o no una alianza con Nikkie.

La estrategia que utiliza el autor va al centro de la cognición orientada a futuro, planteando la siguiente cuestión. Lo que resulta fundamental para decidir el asunto es determinar si los chimpancés tienen un deseo de obtener un rango mayor, y de ser así, si acaso generan conductas positivas para este fin. En definitiva, de lo que se trata es de evaluar si los chimpancés presentan conductas dirigidas a un objetivo o meta. Hoy sabemos que contamos con evidencia a favor de la presencia de comportamiento dirigido a metas en chimpancés, e incluso se ha evaluado comparativamente con humanos en un mismo diseño experimental (Myowa-Yamakoshi *et al.*, 2012). Razonablemente este no era el escenario a la fecha de publicación del libro escrito por de Waal, y no tuvo más recurso que reflexionar acerca de la naturaleza de la orientación a metas. Así, indicó que podemos decir que una conducta X se orienta a un fin Y si X cesa cuando se ha conseguido Y. Si aplicamos este criterio a la evidencia que hemos trabajado en secciones anteriores, entonces podemos decir que existe comportamiento dirigido a metas en el chimpancé que transporta una herramienta a un sitio de alimento, y luego de consumir o guardar el alimento no hace el traslado de una nueva herramienta si ya no hay comida restante. Lo mismo con la acción de abrir una nuez, solo habría orientación a metas si el chimpancé deja de golpear la nuez una vez que esta ha sido abierta, no insistiendo en darle golpes a los restos.

Para ejemplificar con chimpancés, de Waal recurre a un ejemplo externo a su colonia. Hace mención de Goblin, un chimpancé estudiado por Jane Goodall en la colonia de Gombe Stream. La particularidad de la conducta de Goblin está en que se empeñaba en torear a las hembras adultas hasta que estas se sometían a él. De Waal enfatiza que Goodall utiliza adrede la expresión “hasta”, como sugiriendo que el sometimiento de las hembras fue su propósito desde un principio, y es lo que explica el cese de sus agresiones (de Waal, 2022, 304). Situación similar vivió la colonia de Arnhem, refiere de Waal, cuando observó que en el otoño de 1976 disminuyeron las actitudes hostiles de Nikkie hacia las hembras,

justo cuando las hembras comenzaron a saludarlo frecuentemente. El autor insiste en que no cabe ninguna duda que los animales que él observó en la colonia se esforzaban por conseguir un estatus más elevado (de Waal, 2022, 305).

En cuanto a la posibilidad de que los chimpancés realicen predicciones, el autor recuerda la conducta de Yeroen y su insistencia en resistirse a Luit e intentar contacto con Nikkie (de Waal, 2022, 308). Estas interacciones resultaban en experiencias negativas para Yeroen toda vez que no lograba vencer. Su insistencia a pesar de la falta de resultados intenta ser explicada desde una capacidad por la cual Yeroen pudiera predecir el resultado final (como si supiera que “se perdió la batalla, pero no la guerra”).

Por último, el tratamiento que da el autor a la planificación como proceso cognitivo es similar al que hemos dado en este ensayo. La caracteriza como una ruta hacia un objetivo que se compone de sub-metas que permiten conectar la posición inicial del sujeto con una meta lejana (de Waal, 2022, 308 ss.). Para ejemplificar esto recupera dos casos que pudo observar en la colonia. El primero es el de Franje, una hembra que frente al clima frío que se aproximaba trasladó toda la paja que mantenía en su habitación para poder hacer un nido cómodo y temperado en la instalación exterior. Aquí vemos una relación de submetas a meta entre el traslado de los elementos para la construcción del nido y el armado de este. Lo realmente fascinante de este caso es que el traslado ocurre antes de que Franje pueda experimentar el cambio en las temperaturas.

El último ejemplo de planificación en la colonia está en los comportamientos de despedida. La despedida difiere del saludo en que el saludo se realiza usualmente frente a cualquier otro sujeto por el solo hecho de reconocer quien es. La despedida en cambio ocurre cuando se anticipa una separación. El autor pudo observar una peculiaridad llevada a cabo por la chimpancé llamada Gorila, que tiene a su cuidado a la bebé Roosje y debe dar el biberón diariamente. Para esto, los cuidadores la trasladan a la misma hora junto a Roosje a la instalación interior para que pueda alimentarla sin interrupciones. Lo que pudo ser observado fue que Gorila no entraba inmediatamente, sino que se dirigía a donde se encontrara Yeroen y Mama, con el fin de poder tocarlos por unos instantes o darles besos. Si esto no era posible de inmediato Gorila incluso daba rodeos hasta lograrlo. Una vez concluida la conducta se apartaba para ir a las instalaciones interiores (de Waal, 2022, 309). La conducta de Gorila es interpretada como que realiza comportamientos de despedida porque sabe que será separada de ellos en ese momento.

Con esto concluye el análisis del comportamiento orientado a futuro de los chimpancés de la colonia de Arnhem.

## 5. PREDICCIÓN EN EL CEREBRO

Por último, resta el desafío de unificar los elementos aquí presentados y la distinción trazada entre planificación y *timing*. Esta tarea es merecedora de un trabajo aparte, pero creemos que un intento puede ser esbozado desde la ya mencionada hipótesis del procesamiento predictivo.

Se ha propuesto recientemente que las habilidades predictivas complejas que poseemos humanos (y especies cercanas en la escala evolutiva) como son la planificación y la imaginación emergieron de forma gradual en la historia evolutiva desde mecanismos más simples de predicción como pueden ser *loops* básicos de predicción y corrección de errores (e.g reflejos motores y del sistema nervioso autónomo) (Pezzulo *et al.*, 2021). La hipótesis es que estos mecanismos se encontraban ya presentes en los cerebros de nuestros ancestros evolutivos y cumplían un rol importante en la solución de problemas de regulación adaptativa (Pezzulo *et al.*, 2015).

En estudios recientes se ha ofrecido evidencia de que tanto el sistema visual como auditivo operarían conforme a los principios de la teoría de la inferencia activa (Chennu *et al.*, 2013; Kok y de Lange, 2014). Conforme a estos principios, el cerebro forma representaciones neuronales construidas desde experiencias anteriores. Estas a su vez operan como un modelo generativo de cómo es que los estímulos del entorno causarían las sensaciones (Barrett y Simmons, 2015).

Esto implica que el cerebro no operaría como tradicionalmente concebido. Con la revolución cognitiva, no solo se dejó atrás el paradigma conductista, sino que se ofreció una explicación de aquello que ocurría al interior de la mente o cerebro, mediante los conceptos de procesamiento de información y computación, con sus orígenes en las teorías de Shannon (Shannon, 1948) y Turing (Turing, 1937) respectivamente. La imagen de la mente como una gran computadora que manipula información generando un output se encuentra al centro de esta concepción.

En este paradigma, el rol de la mente es pasivo. Si concebimos la mente como un computador cargado con un programa para desarrollar una tarea determinada, tenemos que la activación del algoritmo ocurre únicamente cuando ingresa información en el input de la máquina.

La teoría de la inferencia activa viene a desafiar esta imagen del cerebro como un componente pasivo, ofreciendo en su lugar un cerebro que se encuentra constantemente anticipando el input sensorial, siendo implementado como predicciones que se propagan por las regiones corticales. En este trayecto, modulan el disparo de neuronas a la espera de recibir efectivamente input del entorno. Las predicciones del cerebro se

constituyen así en hipótesis acerca del comportamiento del entorno que pueden ser contrastadas con las señales sensoriales entrantes (Barrett y Simmons, 2015).

La tarea del cerebro es entonces reducir al mínimo la diferencia entre las predicciones y las sensaciones entrantes, esto es, el “error en la predicción”. La diferencia entre el modelo generado y el mundo puede resolverse mediante los sistemas involucrados, esto es, percepción y acción (o control motor) (Parr *et al.*, 2022). Estas dos formas de reducir el error en la predicción se corresponden con las clásicas direcciones de ajuste mente-mundo/mundo-mente (Anscombe, 2000; Searle, 1983).

En términos de implementación, el ajuste ocurre en el cuerpo de tres maneras: (1) propagando el error de regreso por las conexiones corticales para modificar la predicción, (2) moviendo el cuerpo para generar las sensaciones esperadas y (3) cambiando cómo el cerebro atiende o cómo recorta el input sensorial (Barrett y Simmons, 2015).

Hipotetizamos que una modificación de la teoría de la inferencia activa permitiría dar cuenta de los fenómenos de timing y planificación de manera unificada. Lo cierto es que este paradigma permite entender cómo una criatura es capaz de anticipar elementos en su input sensorial de manera situada, y en consecuencia habilitar la explotación de *affordances* existentes en el entorno (Pezzulo y Cisek, 2016). Pero los casos estudiados presentan la dificultad adicional de requerir un cierto nivel de distancia respecto del entorno inmediato.

Esto puede ser salvado si tenemos presente que un modelo generativo puede ser dotado de profundidad temporal, habilitando la inferencia prospectiva. Esto permite la anticipación de las consecuencias de acciones o secuencias de acciones en el corto y largo plazo, además de la planificación e imaginación de situaciones novedosas (Pezzulo *et al.*, 2021).

Investigadores han sugerido que un factor relevante para el desarrollo evolutivo de modelos temporales profundos fue la búsqueda de alimentos; Y que las exigencias cognitivas de la tarea favorecieron el desarrollo de un circuito para la búsqueda retino-telencefálico (procesamiento avanzado) distinto de un circuito retino-tectal (procesamiento temprano) (Cisek, 2019). En esta misma línea, Buzsáki y Moser han propuesto que los mecanismos de memoria y planificación evolucionaron desde mecanismos para la navegación, explotando mapas cognitivos y de predicción ubicados en el hipocampo (Buzsáki y Moser, 2013).

Adicionalmente, se ha planteado que una parcelación progresiva de un modelo inicialmente no-diferenciado del tiempo puede haber sido la forma en que evolutivamente resultaron modelos temporalmente

profundos desde modelos más simples; a su vez, esta parcelación puede haber sido impulsada por la internalización que el animal hace de sus propios patrones de comportamiento secuencial.

## 6. CONCLUSIÓN

En las secciones anteriores desarrollamos la noción de timing en el contexto de la literatura sobre el condicionamiento clásico y operante pudiendo observar cómo la conducta de ciertas especies se modifica en función de componentes temporales de la tarea que están desempeñando, aprendiendo los intervalos relevantes y demorando la respuesta. Distinguimos a continuación la noción de timing de la noción de planificación, uno de los tipos de cognición orientada al futuro más compleja. Logramos apreciar este tipo de cognición a través de ejemplos en la literatura etológica, en concreto por medio del estudio del comportamiento de chimpancés relativo al uso de herramientas y recolección de alimentos. Por último, logramos apreciar diversos tipos de cognición orientada a futuro en el estudio llevado a cabo por Frans de Waal en la colonia de chimpancés de Arnhem, Países Bajos. Cabe notar que, pese a su diversidad, en los casos reseñados podemos atribuir cognición orientada a futuro cada vez que la tarea exige anticipar un evento o estado futuro, sea la obtención de un refuerzo, el traslado a un lugar de destino, el despliegue de una conducta amistosa u hostil, para tender a esta o evitarla.

Existe, no obstante, una distinción que realizar en la atribución de capacidades cognitivas en los casos evaluados. Si bien sostenemos que para el caso de los chimpancés de la selva del oeste africano (quienes planifican su propio desayuno) es posible atribuir planificación que exhibe al mismo tiempo una capacidad análoga al timing, esto es, planificación y timing convergen en ese caso, esto no ocurre en el caso de la colonia de chimpancés en Arnhem, ya que los casos de planificación reseñados no incorporan necesariamente consideraciones acerca de las regularidades temporales de su propio entorno (con quizá la excepción de Gorila, quien experimenta la separación del grupo a diario).

Para el caso de la fuga de "Mamá", la planificación implicó establecer el objetivo de abandonar la instalación y la determinación de pasos intermedios: buscar ramas altas, colocarlas y treparlas. Esto supone una sensibilidad al orden temporal en que deben ocurrir los elementos de una secuencia (se deben colocar las ramas antes de trepar el muro), mas no pareciera involucrar la incorporación de la frecuencia con que ocurre

algún otro evento ni su duración. En el caso de Dandy se establece el objetivo de obtener limpieza y se adopta una estrategia de comunicación gestual para influir en el comportamiento futuro de otros, cuestión que tampoco acarrea la detección de regularidades ni duraciones. Luego, para las conductas de confiscación se anticipa una eventual agresión de un chimpancé a otro y se interviene antes que este resultado se produzca, no requiriendo incorporar intervalos ni duraciones. Por último, para el caso de los enfrentamientos, las conductas para ganar el favor y las conductas de separación apuntan a una capacidad para estimar las consecuencias sociales no inmediatas de sus propias acciones, así como la anticipación de un enfrentamiento próximo, ninguna de las cuales depende de captar duraciones o regularidades temporales. Podemos ver que la planificación, en la medida que involucra la anticipación de un evento o estado futuro constituye genuinamente cognición orientada a futuro, pero dependerá de la tarea específica si el éxito en esta requiere, además, de timing. La mera anticipación de un evento futuro no requiere saber cuándo va a ocurrir, la duración del tiempo intermedio hasta su ocurrencia, ni la frecuencia general con que tiene lugar.

Finalmente, es importante considerar que la evidencia utilizada en la sección sobre timing responde a una metodología científica en la que la observación detallada en entornos controlados prima por sobre la atribución de estados mentales y etiquetas que en otro contexto serían fácilmente atribuibles a humanos, como sí hace de Waal al referirse a los chimpancés de la colonia como teniendo creencias, deseos, metas, emociones, entre otros. Creemos que, al integrar estos estudios experimentales con la literatura propia de la etología, además de los reportes presentes en el trabajo de de Waal, podemos justificar la atribución de cognición orientada a futuro a las especies animales aquí tratadas.

## REFERENCIAS

- Allen, C. (2018). Associative learning. En K. Andrews y J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 401-408). London: Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9781315742250-39>
- Allen-Hermanson, S. (2005). Morgan's Canon Revisited. *Philosophy of Science*, 72(4), 608-631. <https://doi.org/10.1086/505187>
- Anscombe, G. E. M. (2000). *Intention*. Cambridge, Massachusetts: Harvard University Press.



- Barrett, L. F. y Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16(7), Article 7. <https://doi.org/10.1038/nrn3950>
- Boesch, C. y Boesch, H. (1984). Mental map in wild chimpanzees: An analysis of hammer transports for nut cracking. *Primates*, 25(2), 160-170. <https://doi.org/10.1007/BF02382388>
- Bubić, A. y Abraham, A. (2014). Neurocognitive bases of future oriented cognition. *Review of Psychology*, 21, 3-15.
- Bubic, A., von Cramon, D. Y. y Schubotz, R. I. (2010). Prediction, Cognition and the Brain. *Frontiers in Human Neuroscience*, 4, 25. <https://doi.org/10.3389/fnhum.2010.00025>
- Buzsáki, G. y Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience*, 16(2), Article 2. <https://doi.org/10.1038/nn.3304>
- Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., Owen, A. M. y Bekinschtein, T. A. (2013). Expectation and Attention in Hierarchical Auditory Prediction. *The Journal of Neuroscience*, 33(27), 11194-11205. <https://doi.org/10.1523/JNEUROSCI.0114-13.2013>
- Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. *Attention, Perception, & Psychophysics*, 81(7), 2265-2287. <https://doi.org/10.3758/s13414-019-01760-1>
- de Waal, F. (2022). *La política de los chimpancés*. Madrid: Alianza Editorial.
- Dennett, D. C. (1992). *Consciousness Explained*. Boston, Massachusetts: Back Bay Books.
- Dews, P. (1978). Studies on responding under fixed-interval schedules of reinforcement: II. The scalloped pattern of the cumulative record. *Journal of the Experimental Analysis of Behavior*, 29(1), 67-75. <https://doi.org/10.1901/JEAB.1978.29-67>
- Escobar, M., Suits, W. T., Rahn, E. J. y Arcediano, F. (2015). Do long delay conditioned stimuli develop inhibitory properties? *Frontiers in Psychology*, 6, 1606. <https://doi.org/10.3389/fpsyg.2015.01606>
- Froufe Torres, M. (2011). *Psicología del aprendizaje. Principios y aplicaciones conductuales*. Madrid: Editorial Paraninfo.
- Gallistel, C. R. y Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, 107(2), 289-344. <https://doi.org/10.1037/0033-295X.107.2.289>
- Gärdenfors, P. (1995). Cued and detached representations in animal cognition. *Behavioural Processes*, 35(1), 263-273. [https://doi.org/10.1016/0376-6357\(95\)00043-7](https://doi.org/10.1016/0376-6357(95)00043-7)

- Goodall, J. (1964). Tool-Using and Aimed Throwing in a Community of Free-Living Chimpanzees. *Nature*, 201(4926), Article 4926. <https://doi.org/10.1038/2011264a0>
- Gwinner, E. (1986). *Circannual Rhythms: Endogenous Annual Clocks in the Organization of Seasonal Processes*. Cham: Springer-Verlag. <https://doi.org/10.1007/978-3-642-82870-6>
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6(6), 242-247. [https://doi.org/10.1016/S1364-6613\(02\)01913-7](https://doi.org/10.1016/S1364-6613(02)01913-7)
- Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain Research*, 1428, 71-79. <https://doi.org/10.1016/j.brainres.2011.06.026>
- Holland, O. y Goodman, R. B. (2003). Robots with Internal Models: A Route to Machine Consciousness? *Journal of Consciousness Studies*, 10 (4-5), 77-109.
- Janmaat, K. R. L., Polansky, L., Ban, S. D. y Boesch, C. (2014). Wild chimpanzees plan their breakfast time, type, and location. *Proceedings of the National Academy of Sciences*, 111(46), 16343-16348. <https://doi.org/10.1073/pnas.1407524111>
- Kabadayi, C. y Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science*, 357(6347), 202-204. <https://doi.org/10.1126/science.aam8138>
- Kok, P. y de Lange, F. P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology: CB*, 24(13), 1531-1535. <https://doi.org/10.1016/j.cub.2014.05.042>
- Lejeune, H. y Wearden, J. H. (1991). The comparative psychology of fixed-interval responding: Some quantitative analyses. *Learning and Motivation*, 22(1-2), 84-111. [https://doi.org/10.1016/0023-9690\(91\)90018-4](https://doi.org/10.1016/0023-9690(91)90018-4)
- Lowe, C. F., Harzem, P. y Spencer, P. T. (1979). Temporal control of behavior and the power law. *Journal of the Experimental Analysis of Behavior*, 31(3), 333-343. <https://doi.org/10.1901/jeab.1979.31-333>
- McClung, J. S., Goetschi, F., Baumeyer, A. y Zuberbühler, K. (2020). Chimpanzee immigration: Complex social strategies differ between zoo-based and wild animals. *Journal of Zoo and Aquarium Research*, 8(1), Article 1. <https://doi.org/10.19227/jzar.v8i1.326>
- McEwen, E. S., Warren, E., Tenpas, S., Jones, B., Durdevic, K., Rapport Munro, E. y Call, J. (2022). Primate cognition in zoos: Reviewing the impact of zoo-based research over 15 years. *American Journal of Primatology*, 84(10), e23369. <https://doi.org/10.1002/ajp.23369>

- McLean, A. (2001). Cognitive abilities - The result of selective pressures on food acquisition? *Applied Animal Behaviour Science*, 71, 241-258. [https://doi.org/10.1016/S0168-1591\(00\)00181-7](https://doi.org/10.1016/S0168-1591(00)00181-7)
- Milton, K. (1981). Distribution Patterns of Tropical Plant Foods as an Evolutionary Stimulus to Primate Mental Development. *American Anthropologist*, 83(3), 534-548. <https://doi.org/10.1525/aa.1981.83.3.02a00020>
- Morgan, C. L. (1894). An introduction to comparative psychology. London: Walter Scott. <https://doi.org/10.1037/11344-000>
- Myowa-Yamakoshi, M., Scola, C. y Hirata, S. (2012). Humans and chimpanzees attend differently to goal-directed actions. *Nature Communications*, 3(1), Article 1. <https://doi.org/10.1038/ncomms1695>
- Panda, S., Hogenesch, J. B. y Kay, S. A. (2002). Circadian rhythms from flies to human. *Nature*, 417(6886), 329-335. <https://doi.org/10.1038/417329a>
- Parr, T., Pezzulo, G. y Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. Cambridge, Massachusetts: The MIT Press. <https://doi.org/10.7551/mitpress/12441.001.0001>
- Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (pp. xv, 430). Oxford: Oxford University Press.
- Pezzulo, G. y Cisek, P. (2016). Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. *Trends in Cognitive Sciences*, 20(6), 414-424. <https://doi.org/10.1016/j.tics.2016.03.013>
- Pezzulo, G., Parr, T. y Friston, K. (2021). The evolution of brain architectures for predictive coding and active inference. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1844), 20200531. <https://doi.org/10.1098/rstb.2020.0531>
- Pezzulo, G., Rigoli, F. y Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, 134, 17-35. <https://doi.org/10.1016/j.pneurobio.2015.09.001>
- Phillips, C. y Lancelotti, C. (2014). Chimpanzee diet: Phytolith analysis of feces. *American Journal of Primatology*, 76(8), 757-773. <https://doi.org/10.1002/ajp.22267>
- Rao, R. P. N. y Jiang, L. P. (2022, November 22). Predictive Coding Theories of Cortical Function. *Oxford Research Encyclopedia of Neuroscience*. <https://doi.org/10.1093/acrefore/9780190264086.013.328>
- Rescorla, R. A. (1967). Inhibition of delay in Pavlovian fear conditioning. *Journal of Comparative and Physiological Psychology*, 64(1), 114-120. <https://doi.org/10.1037/h0024810>

- Romanes, G. J. (1888). *Animal Intelligence*. Boston, Massachusetts: D. Appleton.
- Sanz, C. M., Call, J. y Boesch, C. (Eds.) (2013). *Tool Use in Animals: Cognition and Ecology*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511894800>
- Sanz, C. M. y Morgan, D. B. (2009). Flexible and Persistent Tool-using Strategies in Honey-gathering by Wild Chimpanzees. *International Journal of Primatology*, 30(3), 411-427. <https://doi.org/10.1007/s10764-009-9350-5>
- Sanz, C., Morgan, D. y Gulick, S. (2004). New Insights into Chimpanzees, Tools, and Termites from the Congo Basin. *The American Naturalist*, 164(5), 567-581. <https://doi.org/10.1086/424803>
- Schneider, B. (1969). A two-state analysis of fixed- interval responding in pigeons. *Journal of the Experimental Analysis of Behavior*, 12, 677-687. <https://doi.org/10.1901/jeab.1969.12-677>
- Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139173452>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sherman, J. G. (1959). The temporal distribution of responses on fixed interval schedules. *Dissertation Abstracts*, 20,1455-1456.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis* (p. 457). New York: Appleton-Century.
- Starzak, T. (2017). Interpretations without justification: A general argument against Morgan's Canon. *Synthese*, 194(5), 1681-1701. <https://doi.org/10.1007/s11229-016-1013-4>
- Zspunar, K. K., Spreng, R. N. y Schacter, D. L. (2014). A taxonomy of prospective: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences of the United States of America*, 111(52), 18414-18421. <https://doi.org/10.1073/pnas.1417144111>
- Thorndike, E. L. (1905). *The elements of psychology* (pp. xx, 353). New York: A G Seiler. <https://doi.org/10.1037/10881-000>
- Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1), 230-265. <https://doi.org/10.1112/plms/s2-42.1.230>
- Vasconcelos, M., de Carvalho, M. P. y Machado, A. (2017). Timing in animals: From the natural environment to the laboratory, from data to models. En J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon y

- T. Zentall (Eds.), *APA handbook of comparative psychology: Perception, learning, and cognition* (Vol. 2, pp. 509-534). Washington: American Psychological Association. <https://doi.org/10.1037/0000012-023>
- Wearden, J. (2016). *The Psychology of Time Perception*. London: Palgrave Macmillan. <https://doi.org/10.1057/978-1-137-40883-9>
- Whitaker, S., Lowe, C. F. y Wearden, J. H. (2003). Multiple-interval timing in rats: Performance on two-valued mixed fixed-interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, 29(4), 277-291. <https://doi.org/10.1037/0097-7403.29.4.277>
- Zieliński, K., Werka, T., Wilson, W. J. y Nikolaev, E. (1995). Inhibition of delay of the two-way avoidance response and warning-signal salience. *Animal Learning & Behavior*, 23(4), 438-446. <https://doi.org/10.3758/BF03198943>



eISSN: 1989-3612  
DOI: <https://doi.org/10.14201/art2024.31142>

# THEORY-CONSTRUCTION IN COMPARATIVE COGNITION: ASSESSING THE CASE OF ANIMAL NORMATIVITY

## *Construcción de teoría en cognición comparada: evaluando el caso de la normatividad animal*

Nicolás SÁNCHEZ  
*Universidad Nacional de Córdoba, Argentina*  
nssanchez.unc@gmail.com

Recibido: 29/12/2022      Revisado: 12/08/2023      Aceptado: 09/11/2023

**ABSTRACT:** With an extensive amount of research on the social lives of primates, Frans de Waal has been a pioneering advocate for the continuity of human and non-human minds, putting forward the idea that these creatures exhibit rudimentary political and moral behaviors. One of the traits which de Waal focuses on is *animal normativity*, a set of behaviors functionally defined as adherence to social standards. Recently, some philosophers have endorsed this position, holding that animals show a psychological capacity called *normative cognition* underlying those and other social behaviors. In this paper, I assess whether advocacy for animal normativity is an exercise of theory construction in comparative cognition. To that end, I present three features of this kind of theory construction. First, the explanatory goal of building functional analyses of cognitive capacities. Second, the conceptual aid of comparative thinking for theory construction. Third, the heuristic value of theory in specifying possible roads of inquiry. Taking these features into account,

I assess whether the claims advocates make regarding animal normativity consider them. My answer is negative. First, since some advocates focus on behavioral traits and not on psychological capacities, they are not producing theory in comparative cognition, although, as I argue, they should. Second, there is a disregard for hypothesis testing and no evolutionary considerations to support their views. Finally, the claim that non-human animals exhibit normativity does not seem to have heuristic value.

*Keywords:* animal minds, comparative psychology, normative cognition, functional analysis, evolution.

**RESUMEN:** A partir de una gran cantidad de investigación respecto a las vidas sociales de los primates, Frans de Waal ha sido un pionero defensor de la continuidad mental entre animales humanos y no humanos, avanzando la idea de que tales criaturas exhibían rudimentos de comportamientos políticos y morales. Uno de los rasgos en el que se ha concentrado Frans de Waal es el de *normatividad animal*, un conjunto de comportamientos funcionalmente definidos como la adherencia a estándares socialmente. Recientemente, a su vez, algunos filósofos y filósofas han apoyado esta posición, afirmando que los animales muestran una capacidad psicológica de *cognición normativa* que subyace a aquellos y a otros comportamientos. En este artículo, evalúo si la defensa de la normatividad animal constituye un ejercicio de construcción de teoría en cognición comparada. Con ese fin, presento tres rasgos de esta clase de construcción teórica. En primer lugar, el rol explicativo de construir análisis funcionales de capacidades cognitivas. En segundo lugar, la ayuda conceptual que brinda el pensamiento comparativo a la cognición comparada. En tercer lugar, el valor heurístico de la teoría en especificar caminos posibles de indagación. Tomando en cuenta estos rasgos, evalúo si las afirmaciones de los defensores de la normatividad animal los toman en consideración. Mi respuesta es negativa. En primer lugar, dado que algunos defensores se concentran en rasgos comportamentales y no en capacidades psicológicas, no están produciendo teoría en cognición comparada, si bien, como argumento, deberían. En segundo lugar, hay una despreocupación por el testeo de hipótesis y no hay consideraciones evolutivas que apoyen la posición de los defensores. Finalmente, la afirmación de que los animales no humanos exhiben normatividad no parece tener un valor heurístico definido.

*Palabras clave:* mentes animales, psicología comparada, cognición normativa, análisis funcional, evolución.



## 1. INTRODUCTION

Research on animal cognition and the evolutionary origins of cognitive capacities put pressure on traditional claims about human uniqueness. For decades, Frans de Waal has been a pioneering advocate for the continuity of human and non-human minds, a position he has put forward in scientific work (de Waal, 1989) and writings directed to a wider audience (de Waal, 2016). His extensive research on the social lives of primates supports the idea that these creatures exhibit rudimentary political and moral behaviors, questioning the status of human beings as the only moral and political animals. Some of his findings concern quantitative analyses of post-conflict resolution (de Waal, 2007) and the design of experimental tasks -the inequity aversion behavioral paradigm (Brosnan and de Waal, 2003). Other contributions concern claims about research on non-human animals -e.g., his claim that anthropodenial is as dangerous as anthropomorphism (de Waal *et al.*, 2006).

One of the traits on which de Waal focuses is *animal normativity*, a set of behaviors functionally defined as adherence to social standards. These behaviors are evidenced by respecting social hierarchy or reconciliatory behaviors that primates exhibit to maintain and restore social harmony. Recently, some philosophers have endorsed de Waal's position, holding that animals show a psychological capacity called *normative cognition* -or ought thought- underlying those and other social behaviors. This psychological trait implies identifying and internalizing social patterns, tracking norm-compliance, and punishing non-conformers.

In this paper, I will assess whether advocacy for animal normativity is an exercise of theory construction in comparative cognition. To that end, I will first present the advocates' approach, focusing on *which* behaviors they take as evidence of animal normativity and *why* those behaviors are so considered (section 2). Then, I will present three core features of theory construction in comparative cognition (section 3). The first feature concerns the explanatory goal of building functional analyses of cognitive capacities. Second, the conceptual aid of comparative thinking for theory construction -a style of reasoning that helps to get an evolutionary profile of the creatures under analysis. Third, the heuristic value of theory in specifying possible roads of inquiry.

Finally, I will assess whether what advocates claim regarding animal normativity present the core features of theory construction in comparative cognition (section 4). My answer will be negative. First, since some advocates -like de Waal- focus on behavioral traits, they are not producing

theory in comparative cognition. Furthermore, in light of recent arguments claiming that animal normativity could work at the behavioral level (Westra and Andrews, 2022), I will argue why advocates should care about psychological capacities. Second, as long as advocates focus on psychological traits, my argument will be that their claims do not exhibit the core features of theory construction in comparative cognition. In particular, there is a disregard for hypothesis testing and no evolutionary considerations to support their views. Finally, the claim that non-human animals exhibit normativity does not appear to have heuristic value.

On a more general note, while my present concern regards the specific case of animal normativity, a broader purpose is to provide -the sketch of- a framework to assess what intellectual efforts count as contributions to the field of comparative cognition. Given that some philosophers attempt to contribute to the theory of animal cognition, this -sketch of a- framework could be valuable as a guide.

## 2. ADVOCATES FOR ANIMAL NORMATIVITY

Identifying social patterns of behavior, internalizing them, sanctioning and tracking norm-compliance are deeply entrenched features of what makes us human, so much so that concepts like ‘norm’ and ‘normativity’ “occur in an enormous range of research that spans the humanities and behavioral sciences” (Kelly and Setman, 2021). Furthermore, this “capacity for understanding the difference between how things are and how things ought to be” (Schlingloff and Moore, 2017, 381) is considered to be unique to humans, entitling us to the coveted status of *normative animals* (Roughley and Bayertz, 2019).

Although most research on norms in the social sciences has focused on adult humans, more recent efforts have inquired about the phylogenetic and ontogenetic origins of these capacities. Phylogenetically, researchers hypothesize that challenges regarding the evolution of cooperation in Pleistocene societies –the need for delayed returns cooperation- selected for social norms in the hominin lineage (Sterelny, 2019; 2021). Ontogenetically, children exhibit normative attitudes in social interactions -enforcing norms as third parties when others violate them in their presence- at approximately three years old (Schmidt and Rakoczy, 2019). Crucially, that capacity emerges developmentally with other social skills -e.g. over imitation.

Questions about the phylogenetic origins of normative capacities have also been pressed in the animal world, especially in trying to account for

the origins of morality<sup>1</sup>. As with most human-like features, different attitudes permeate the approach to the answers, and a lot depends on how the target phenomenon is defined. Skeptics claim that the gulf between humans and non-humans is too deep to call normative what animals do (Schmidt and Rakoczy, 2019). However, advocates for animal normativity claim that “some non-human animals can actively respond to norms and even have some understanding of them” (Danón, 2019, 177). In this sense, empirical evidence supports “the idea that [...] normativity extends beyond the boundaries of the human species” (Monsó and Andrews, 2022, 209; for similar claims, see Andrews, 2020, 52; Fitzpatrick, 2020, 45).

Frans de Waal is one such advocate. In his studies regarding the social behavior of primates, he construes behaviors like reconciliation or respect for social hierarchies as “adherence to an ideal or standard” (de Waal, 2014, 287). Reconciliation is a friendly post-conflict behavior that occurs no more than half an hour after a fight in which one of the parties -irrespective of dominance- approaches the other and makes begging gestures like stretching an arm with an open hand. These encounters involve substantive eye contact and kissing (de Waal, 1989, 42). The function of reconciliation is to maintain social harmony, with the evolutionary-oriented rationale of

restor[ing] relationships that have been damaged by aggression but are nonetheless essential to survival. Since many animals establish cooperative relationships within which conflict occasionally arises, mechanisms of repair are essential (de Waal, 2014, 192).

These -and other- behaviors comprise a functional kind that de Waal calls animal normativity.

1. The philosophical significance of the phylogenetic origins of normative cognition is not exhausted by the emergence of morality. Empirical findings and theoretical reflections in this domain can illuminate two other subjects that interest philosophers. First, the conceptual (or metaphysical) claim that thought is normative. Understood as the claim that what a creature thinks is determined by the rules they are capable of following, findings about the origins of normative cognition would have consequences for theories about the origins of thought. Second, the philosophical use of this capacity with demarcational purposes -i.e., as a measure of cognitive complexity. According to this view, a creature endowed with normative cognition is more cognitively sophisticated -in some sense- than a creature without it. Though relevant and influential, these issues are not the focus of the present paper. My sole focus here concerns the origins of particular behaviors and psychological capacities. In that regard, the discussion is independent of claims about the nature of thought or about kinds of minds.

Building upon de Waal's pioneering work, a group of philosophers also construe some behaviors as evidence for normative capacities (Fitzpatrick, 2020; Andrews, 2020; Vincent, Ring, and Andrews, 2019). Two shifts in focus in this latter approach are worth mentioning. First, for these philosophers, the question no longer concerns morality but rather "the broader issue of whether some animals have a general capacity to act in the light of norms" (Danón, 2019, 177). Second, the focus is on a psychological rather than a behavioral trait. While de Waal held vaguely that a "preference for certain social outcomes over others" (de Waal, 2014, 201) guided normative behaviors, these latter advocates make specific claims regarding the psychological capacities that underpin them.

Vincent, Ring and Andrews, for example, claim that research about normative behavior should focus on a

basic cognitive requirement for moral agency—namely, ought-thought, which is a cognitive modality much like mental time travel or counterfactual thinking. Thinking about what ought to be the case—like thinking about what happened in the past, what might happen in the future, and what might be the case under various circumstances—is a cognitive mode that requires the thinker to do more than represent what is currently the case. (Vincent, Ring, and Andrews, 2019, 58-59)

Ought-thoughts, in this sense, are more general than moral oughts because they are applied more widely through diverse "instances of valuing, some of which are not moral" (Vincent, Ring, and Andrews 2019: 59). This capacity is expressed through normative practices, defined as "patterns of behavior shared by members of a community that demonstrate they value certain ways of doing things as opposed to others" (Vincent, Ring, and Andrews, 2019, 59).

The strategy for advocates is to look for behaviors -expanding over de Waal's list- that qualify as normative practices. Those behaviors are, for example, cultural differences between groups of chimpanzees. Different cultural traditions between groups, like tool selection for nut cracking -i.e., some groups prefer wooden tools and some groups tools made of stone (Luncz and Boesch, 2014)- indicate preferences for doing things inside the group. When females from one group migrate to another, the ranking of those females depends on their learning the new group's tradition. Another example is the alleged norms the chimpanzees follow when they hunt colobus monkeys (Vincent, Ring, and Andrews, 2019, 59). This behavior is described as a "coordinated division of labor, with some individuals driving monkeys in the direction of groupmates waiting to ambush" (Fitzpatrick,

2020, 5). Given that the hunt's rewards are typically distributed "in proportion to each individual's level of participation" (Fitzpatrick, 2020, 5), a straightforward interpretation of this behavior is that there are norms of cooperation that regulate this practice.

Furthermore, these behaviors count as normative practices because of their resemblance to human activities:

We recognize certain patterns of behavior in humans as constituting certain phenomenon (e.g., friendship, policing, babysitting, etc.). When we see similar patterns effecting similar ends in other social species, we are warranted in classifying them as instances of these same phenomena. That is what consistency requires, so long as the attribution of friendship, policing, or babysitting still seems apt after all the known and likely capacities of the species in question have been taken into account (e.g., their capacities for emotion, social learning, creative problem solving, etc.). We need to develop or embrace definitions of the phenomena in question to allow for unambiguous attribution to both humans and other animals. For instance, if we define 'friendship' in terms of developing and maintaining affiliative social bonds, then it would be an error to deny that chimpanzees have norms of friendship (Vincent, Ring, and Andrews, 2019, 75).

In a single-authored paper about ought-thought, Andrews pursues a different strategy. Taking Bicchieri's definition (Bicchieri, 2016) as a guide, Andrews identifies three conditions for social norms:

- (1) there is a pattern of behavior demonstrated by community members;
- 2) individuals choose to conform to the pattern of behavior; and
- (3) individuals expect that community members will also conform and will sanction those who do not conform (Andrews, 2020, 40).

Furthermore, primates need these four abilities to meet these conditions: being able to identify group members as agents, distinguish the behavioral patterns of their group from out-group behavior, learn their group traditions socially, and respond to inappropriate behavior through sanctions. Primates would exhibit all of these cognitive requisites: they ascribe goals to other agents, are sensitive to in-group out-group differences -as evidenced by the example of female immigration-, learn traditions socially -as evidenced by the example of cooperation and, again, female immigration- and punish non-conformers through aggressive behavior. For Andrews, all this counts as evidence of ought-thought.

Having described the approach to animal normativity from advocates, I will present the features of theory construction in comparative cognition.

### 3. THEORY CONSTRUCTION IN COMPARATIVE COGNITION

Two styles of reasoning are recognized as the main ones to construct theories in comparative cognition (see Papini, 2021; Vonk and Shackelford, 2012; Wasserman and Zentall, 2009)<sup>2</sup>: functional analysis and the comparative approach.

The first style of reasoning is functional analysis, that consists in

analyzing a disposition into a number of less problematic dispositions such that programmed manifestation of these analyzing dispositions amounts to a manifestation of the analyzed disposition (Cummins, 2010, 291).

These analyses are functional because their components are selected by virtue of what they “do [...] rather than in terms of their intrinsic constitutions” (Roth and Cummins, 2017, 35).

To get functional analyses, scientists need to start with behavioral *phenomena* or *effects*, behavioral regularities or dispositions of the system that needs explaining, found through behavioral tasks or paradigms. The performance of New Caledonian crows in the Aesop Fable paradigm is, for example, one of those effects. In this task, scientists present crows with a water-filled tube containing a reward floating on a cork. Under these circumstances, crows push stones into the water until they reach the food (Bird and Emery, 2009). The crows’ systematic performance in the task is the behavioral effect.

Having found effects, scientists provide a speculative answer as to what psychological capacity the system might be exercising. In the case of the Aesop Fable task, the hypothesized trait is causal understanding. After several manipulations of the conditions of the test, researchers hold that

[t]his paradigm can therefore be used as a test of causal cognition, investigating whether animals can understand or learn about various causal regularities which underlie the displacement of water (Jelbert *et al.*, 2014, 2).

Ideally, further research will analyze that capacity into sub-capacities. However, that level of detail is rare in comparative cognition, and most discussions stay at the level of traits that underlie behavioral performance.

2. Under some views that define cognition narrowly, only a subset of capacities studied by comparative psychology concern cognition (see Olmstead and Kuhlmeier, 2022). In what follows, I take it as not controversial that comparative psychology does study cognition -focusing on proximate mechanisms- through a specific focus concerning phylogeny and adaptation.

This approach to behavioral tasks and to assessing what capacity it is that they measure are consistent with what Shettleworth considers an indication of well-designed research in comparative cognition: “progress [in comparative psychology] is greater if the process under study is well defined for [the species]” (Shettleworth, 2009, 213). To be well-defined is to have a behavioral task that is relatively clear on the psychological capacity it tests. In that sense, scientists can use the Aesop Fable task to test whether a given species exhibits causal understanding.

The second style of reasoning relevant to comparative cognition is what Currie (Currie, 2021) labels comparative thinking -an approach applied widely through the spectrum of the biological sciences. The purpose of comparing traits in biology is to construct an evolutionary profile of the creature under analysis. To this aim, two processes are especially relevant: processes that “act as connecting lineages” (Currie, 2021, 5) - like sex- that create “causal and ancestral connections” (Currie, 2021, 5), and processes “creating variation within and between lineages” (Currie, 2021, 5) - like adaptation to environmental pressure. These two processes explain the similarities biologists find between traits of different species. The two most common similarities found in biology are homology and analogy, similarities explained through common ancestry and adaptation to environmental pressure, respectively.

In comparative cognition, comparisons help in theorizing about cognitive capacities across species. In that sense, Emery and Clayton have argued that convergent evolution explains why primates and crows have similar abilities for causal reasoning and prospective thought, given that both species would face

the same socioecological challenges, such as locating perishable food distributed in time and space or understanding the relationships between different individuals within large social groups. (Emery and Clayton, 2004, 1905)

Another influence of comparative thinking in comparative cognition comes from the rationale it provides to Morgan’s Canon, an application of Occam’s razor that instructs scientists to choose the simpler of two competing psychological hypotheses. Critics of this parsimony principle have pointed out that the conceptual problem is to cash out simplicity and complexity in this context (for critical discussion, see Buckner, 2013; Fitzpatrick, 2017). Comparative thinking provides a rationale for that specification: simpler processes are those more widely distributed phylogenetically. Thus, when comparative psychologists have two hypotheses –concerning basic learning processes and higher cognitive traits, for

example- to account for the same phenomena they should choose the more plausible, the more widely distributed phylogenetically. The suggestion is to read the principle not functionally -stating that simpler processes allow creatures to do 'less' than more complex ones- but evolutionarily: widely distributed traits versus particular adaptations of particular species (Currie, 2021, 38)<sup>3</sup>.

The third and final core feature of theory construction in comparative cognition concerns its heuristic value. Take, for example, Povinelli's reconstruction of Gallup's theory regarding the mirror test. For Gallup, adequate performance in this test implicates some capacity for self-awareness, and thus it predicts a capacity for introspection and social attribution (Povinelli, 1993, 497). Although that model has received several objections regarding the necessary connections between performance in the test and the underlying psychological processes mentioned above, Povinelli emphasizes that its virtue lies in "being a way of generating strong (and hence quite falsifiable) predictions" (Povinelli, 1993, 498). Evolutionary considerations can also provide this heuristic value. Function -understood broadly as ultimate or proximal (Tinbergen, 1963)- can direct questions about proximal mechanisms and, in that sense, guide specific empirical research. Thus, biological theories that assert that kinship and relatedness between organisms mediate the evolution of social behavior raise questions about how recognizing something as kin or a potential mate increases reproductive success (Sherry, 2005, 444).

From this characterization of comparative cognition as a way of studying cognition aided and permeated by comparisons, we can extract three core features of theory construction in this field:

1. Comparative cognition explains psychological capacities through functional analysis.
2. Comparative thinking aids theory construction in comparative cognition.
3. Theory in comparative cognition has a fundamental heuristic value.

3. Morgan's original formulation of the canon already suggested that we should construe "lower" and "higher" in evolutionary terms. The now standard interpretation of the principle as an application of Occam's razor is a consequence of its reception by psychologists of behavioristic leanings (see Sober, 2005; especially fn. 9), a reception that jettisons the evolutionary content of the principle.



These core features are not all-or-nothing categories but rather gradual ones. However, for a claim to be an exercise in theory construction of comparative cognition, at least the three core features must be present to some degree.

#### 4. ADVOCATES FOR ANIMAL NORMATIVITY MEET COMPARATIVE COGNITION

Having reviewed the claims made by advocates for animal normativity and the core features of theory construction in comparative cognition, I will assess whether the theoretical claims made by advocates present those core features.

##### 4.1. *The scope of the claim about animal normativity: behavioral or psychological trait?*

As mentioned above, some advocates of animal normativity focus on behavioral traits. Along the same line, Westra and Andrews (2022) hold that normativity can be identified as a behavioral phenomenon while remaining non-committal regarding psychological capacities. As I mentioned in section 3, one of the features of theory construction in comparative cognition is to make claims about psychological capacities. Thus, given that my present aim is to assess whether advocates make theoretical claims in comparative cognition, this position could be dismissed for not having a core feature of this domain. However, that would be too hasty. I want to explore next why this position would be unattractive to the advocates' own aims.

Two questions are especially relevant in this context. First, what kind of information is provided by the claim 'animals exhibit normativity', if we understand normativity as a behavioral phenomenon? Second, what kind of inferences are we entitled to draw from it? Regarding the first question, normative behavior -as presented by advocates- appears to be a functional construct comprising different particular behaviors. It is not a single trait but rather a set of them. When studying animal behavior, behaviors that belong to the same kind make a *domain*. Ethology, behavioral ecology, and comparative psychology organize their fields through these functional kinds -e.g., foraging, tool use, or spatial navigation. These disciplines describe domains *functionally*, in terms of what a creature can do -to get food, to use tools, or to navigate space. This taxonomic

organization allows for comparisons between different species regarding how they achieve the same thing. Normative behavior could be, under this approach, a domain: a functional kind defined as “an ability to internalize and enforce social norms” (Fitzpatrick, 2020, 1).

Regarding the second question, and while acknowledging that this theoretical choice is available for advocates, it would come at a cost. The statement “nonhuman animals exhibit normative behavior” would have to be treated very much like the statement “animals exhibit abilities for spatial navigation”. Those statements provide very little empirical information about how the animals do what they do, so it would not allow advocates the inferences about cognitive similarities between normative behavior in humans and normative behavior in animals they seem to want to make. Those inferences about *mental* similarity would not be plausible since different capacities may underwrite those two behaviors -e.g., bees and taxi drivers exhibit abilities for spatial navigation but through diverse cognitive means. However, advocates want to claim that there is a mental continuity in humans and animals concerning normative behavior. It is hard to assess in what way we are continuous<sup>4</sup>.

To put the contrast most starkly, consider an example from behavioral ecology. A sender-receiver model in behavioral ecology describes and predicts communicative interactions between -at least- two systems. Under this framework, both bacteria and I do the same thing: send and receive signals. In a sense, we are both communicating -the abstract functional kind- but to claim some continuity between bacteria and me regarding communication would not be to say much. That is because the sender-receiver model works only at the behavioral level and does not make claims about the internal architecture of the systems. What licenses researchers' and philosophers' inferences about mental continuity are commitments to

4. Another issue in passing: domains can be gerrymandered. Advocates collect behaviors as examples of normativity because those behaviors resemble our own. However, as mentioned above, if comparative thinking does not explain those similarities, their epistemic value of the functional kind is unclear. For example, there was a time when scientists thought they could teach animals how to use language -albeit in a rudimentary way. Handbooks in comparative cognition mirrored this tendency by defining these behaviors functionally as *language-use* (see, for example, the index in Greenberg and Haraway, 1998). This functional categorization of behavior as ‘language-use’ defined that domain through a human perspective. The lesson to extract here is that while comparative thinking constrains the interpretation of functional kinds regarding animal behavior, the absence of those constraints may lead to anthropomorphization. If we focus on the ultimate or evolutionary function of behaviors included by advocates under the tag ‘animal normativity’, we find different phylogenetic origins of those traits. That may suggest that ‘animal normativity’ is a gerrymandered functional kind.

psychological capacities. Thus, advocates for animal normativity *should* see themselves as committing to claims about psychological traits and not only behavioral ones.

#### 4.2. *What guides advocates in their theoretical claims about psychological capacities?*

Restricting the scope of the assessment to theory-construction about psychological capacities, I believe that the core features of comparative cognition are not present. In the following subsections I will make that case.

##### 4.2.1. *Advocates and hypothesis-testing*

As mentioned above, constructing functional analyses in comparative cognition requires testing hypotheses usually<sup>5</sup> through behavioral tests. Advocates claim that there is a specific cognitive capacity called ought thought, but the evidence presented is not precise enough to discard other hypotheses about psychological traits. Crucially, the evidence presented does not exclude the possibility that the creature is acting self-interestedly rather than being aware of what she ought to do. Furthermore, there is no discussion that third-party norm enforcement is a necessary feature of normative cognition, as many researchers claim.

As we mentioned, a piece of methodological advice to see if a creature has a capacity is to apply -or to analyze performance- in the same test that another organism had passed. The advocates' approach does not consider this strategy. Schmidt and Rakoczy (Schmidt and Rakoczy, 2019, 131-134), for example, review the negative experimental evidence that shows that chimpanzees fail to punish as third parties or that apes are not sensitive to fairness in a behavioral version of the ultimatum game paradigm. The absence of discussion of this evidence on the part of advocates seems symptomatic of a disregard for hypothesis-testing.

Besides the lack of discussion of negative evidence, the claim that the kinds of evidence advocates analyze are evidence of ought thought is controversial. Disregarding the operationalization of normative cognition

5. This qualification signals that the relevant contrast is not between experimental behavioral tests and naturalistic observation but between mere description and hypothesis testing (that researchers can also perform in the wild).

for humans, advocates focus on a conceptual definition and assess whether it is suitable for the animal world<sup>6</sup>. By focusing only on the latter but without a behavioral way to measure it, advocates look for superficial similarities between our and the other animals' "normative" behaviors as evidence for the claim that the same cognitive capacity is at work. However, this approach to the conceptual definition would only be viable by considering the relevant behavioral task. Without that constraint, it is easier to bend what counts as 'normative' and to assume that the similarity found is explained through the same cognitive capacity. The problem with this approach is that the inferential step from similar behavior to similar psychological trait is the hypothesis to be *tested*, not to be assumed. Thus, the evidence does not support a specific claim about cognitive capacities.

In the case of Andrews's strategy to find cognitive prerequisites for ought-thought, the evidence receives a particular treatment as she analyzes the evidence for each cognitive requisite separately. For example, evidence of social learning of group traditions is exhibited in patterns like border patrols by chimpanzees. They display punishment, on the other hand, in inequity aversion and prevention of infanticide. However, Bicchieri's definition was that the behavior analyzed exhibited *all* the cognitive markers of normative cognition. Analyzing each behavior in this way gives the impression that primates meet all cognitive requisites. However, what is needed is a single behavior that presents all of them, not four different behaviors -with distinct phylogenetic origins, if I may add- that each satisfies a prerequisite.

#### 4.2.2. Advocates and comparative thinking

As we mentioned, one of the strategies for advocates was to collect behaviors that look alike. However, the idea behind comparative thinking was that similarities could and should be explained through an evolutionary profile of the creatures compared. This feature also seems to be lacking in the advocates' approach.

6. In addition, there seems to be confusion about the meaning of *operationalization*. Vincent *et al.* say that "the operational definition should not demand more than what we typically regarded in the human case" (Vincent, Ring, and Andrews, 2019, 74). But in the next paragraph, arguing this point, they claim that "if we define friendship in terms of developing and maintaining affiliative social bonds, then it would be an error to deny that chimpanzees have norms of friendship" (Vincent, Ring, and Andrews, 2019, 75). But this is a conceptual definition, not an operational one.

On the one hand, it seems that the inferred psychological capacities are not analogous: there is no specification of socioecological problems that may have given rise to ought thoughts. We can find examples of this kind of reasoning regarding social norms in hominins. Tomasello claims that cooperation needed in group hunting requires intersubjective norms to function (Tomasello, 2016). Sterelny, on the other hand (Sterelny, 2019; 2021), claims that norms are a much more recent artifact, necessary only when cooperation becomes much more complex -i.e., delayed returns mutualism: I fix your bow today and expect a part of your hunt in the near future. The question for advocates is what kind of socioecological problem was or is present in the lives of chimpanzees for ought thoughts to emerge. A useful speculative tool in this context is thinking about contrast classes: what was life like when there was no normative cognition? (Currie, 2021, 3). This style of reasoning is not found in advocates. However, an in-principle difficulty for advocates in this sense derives from the diversity of behaviors that show the capacity. Cooperative hunting and respect for social hierarchies seem to solve two different needs for a community -the multiple evolutionary benefits of social stability versus the specific benefits of getting resources. The standards enforced by patrolling borders, furthermore, “likely reflects [...] an adaptive drive to protect scarce resources” (Schlingloff and Moore, 2017, 386). If analogical reasoning is a way of guiding theory regarding psychological capacities, these capacities may be very different -they respond to different needs- and it is not clear why they need to be normative.

Regarding homological reasoning, it is fair to say that it is hard to track cognitive homologies, given that scientists identify most homologies through morphological structure, and psychological capacities do not leave fossil records. However, two traits may share a phylogenetic origin, and that process may allow the inference that they are similar. Advocates do not seem to pay attention to this feature, so they include normative behaviors that seem phylogenetically ancient -like respect for social hierarchy- with more recent acquisitions -social learning of group traditions. Wolves and many mammals whose survival depends on social stability exhibit respect for order. However, the mechanisms underlying these behaviors may be hardwired in a way that patrolling borders and cooperative hunting are not. Since all of these behaviors seem to come at distinct phylogenetic periods, treating them as an exercise in the same cognitive capacity seems far-fetched.

The lack of evolutionary considerations shows again in Andrews’s approach to applying a conceptual definition from humans to the animal world. Andrews modifies Bicchieri’s view, so it does not require

mind-reading abilities. In this modification, we can also see the lack of comparative thinking. There is no justification regarding how there could be one without the other. All we get is

we can ask whether there are human practices that look like social norms but lack the mindreading aspect. (Andrews, 2020, 38)

The inference here seems to be that the same cognitive capacity may be present in a simple or sophisticated manner in animals and humans -with or without mind-reading abilities. What I take to be the problem here is that since this inference is not rooted in comparative thinking, there are no grounds for the speculation that there may be ought thought without mindreading. Some critics claim that such conceptual moves around normativity amount to changing the subject regarding the target phenomena (Schlingloff and Moore, 2017, 381). That may be true, but the deeper problem is that speculation becomes too unconstrained, given that what is simpler and more complex is decided on not-comparative grounds.

#### 4.2.3. Advocates and heuristics

One of the purposes of theory construction in comparative cognition is to provide a basis for future research. For that to work, the idea of researchers in comparative cognition is that their hypotheses have heuristic value only if they are claims of necessity about phenomena. Thus, as long as they have empirical content, these claims allow us to make predictions. Povinelli's interpretation of Gallup's mirror test presented in the previous section is telling in that regard. In the specific case of normative cognition in human infants, it seems to come together developmentally with the capacity for over-imitation (Schmidt and Rakoczy, 2019, 128).

What do we find in animal normativity advocates? Instead of necessity, there is a focus on possibility in the following sense: moral agency does not require capacities for second-order reflection while following social rules does not require mind-reading abilities. The strategy is to identify a demanding characterization of human normative cognition and then offer a more deflated definition in which the phenomenon of interest is still recognizable -based on superficial similarities, as mentioned above. However, how does negating a necessity claim contribute to theory construction? What do we expect to find? The idea may be that normative capacities are not only human but that, in itself, is not of heuristic value.

Theoretical claims with potential heuristic value are not absent from the advocates' approach, although they are not fully examined. Andrews makes specific claims about ought thought as a cognitive mode that allows a creature to think of possible -as opposed to actual- situations. It is a way of thinking about possible states of affairs and reasoning with these alternatives. This cognitive mode would comprise abilities for counterfactual reasoning, prospective planning, and normative cognition. We can expect that performance in behavioral tests for counterfactual reasoning would correlate well with performance in tests for the other abilities<sup>7</sup>.

Theory-construction about animal normativity does not seem to provide a specific research program and thus it does not count as providing heuristic value. If, as de Waal holds, "interpretative labels are only as good as the hypothesis that they generate" (de Waal, 1991, 301), then, it seems that animal normativity, as an interpretative label, is not very good at all.

## 5. CONCLUSIONS

In this paper, I have assessed whether advocacy for animal normativity -a position held by Frans de Waal and several philosophers interested in animal minds- constitutes an exercise of theory-construction in comparative cognition. To make this assessment, I first offered a characterization of the claims made by advocates of animal normativity (section 2) and then the core features of theory-construction in comparative cognition (section 3). The overall argument was negative and it proceeded in steps. First, I assessed the scope of the claim: when advocates say that animals exhibit normativity, are they making claims about *behavior* or *cognition*? At that point, I argued that, if advocates want to establish a cognitive continuity between nonhuman animals and us, they should commit to a claim about psychological capacities and not only to a claim about behavioral traits. Second, I assessed whether advocates advanced their claims by focusing on hypothesis-testing. At that point, I argued that the focus was on conceptual but not operational definitions of normative cognition, an attitude that expressed disregard for hypothesis-testing. Third, I considered whether comparative thinking was a style of reasoning that advocates employed and found that they did not take on homological or

7. Remember, however, the absence of behavioral tasks that measure normative cognition and the negative evidence of existing tests mentioned in 4.2.1.

analogical reasoning. In the first case, advocates seem to lump together behaviors that have distant phylogenetic origins. In the second case, advocates do not specify socioecological challenges that may have selected for normative cognition. Finally, I considered if the claims made by advocates had any heuristic value. In this sense, I argued that advocates were more interested in negating necessity claims regarding human normativity than in establishing necessity claims regarding animal normativity. I argued that claims about animal normativity need the latter. Thus, advocates appear not to be offering hypotheses with heuristic value. The overall argument is that animal normativity as a functional kind is not an exercise of theory construction in comparative cognition.

In this sense, my assessment regarding the functional kind of normativity is consistent with the verdict of researchers that claim that the absence of third-party punishment is a strong indicator of the difference between animals and us (Schmidt and Rakoczy, 2019, 131-134, Sterelny, 2019). Even researchers that are more sympathetic to the idea of continuities between animals and humans are careful to point out that we only find *precursors* of social norms in *specific* behavioral domains rather than stating that animals exhibit normativity (Luncz et al., 2018, 69; van Schaik, 2012, R403). However, I would like to point out two caveats concerning the implications of my criticisms. First, some of my objections to the advocates' position have to do with a lack of evidence that would support the idea that animals' behaviors are normative, mainly because the evidence provided does not discard alternative hypotheses. It is crucial to notice that, as the saying goes, absence of evidence is not evidence of absence and that further empirical research may provide further behavioral tests, evolutionary considerations or claims with heuristic value that makes the advocates' position more plausible. On the other hand, my objections to the advocates' strategy focus on their attempts to establish the general idea that "normativity extends beyond the boundaries of the human species" (Monsó and Andrews, 2022, 209). However, this is a criticism about a particular functional kind that seems to be too vague –it comprises too many behaviors- to specify a single cognitive capacity that shows that normative cognition is also present in nonhuman animals. Advocates could focus on more specific functional kinds to put forward more plausible hypotheses about cognitive continuities.

As a final consideration, I believe the approach followed in this paper may be helpful in assessing other claims about animals and their similarities with humans. The advantages of the view presented here are that it establishes clear guidelines for theory construction -or core features- in



comparative cognition. These guidelines are not advanced to get philosophers to design behavioral tests of normative cognition -or any other psychological capacity. In that sense, my critical remarks do not intend to compel philosophers to start doing experimental science (contrast with Allen, 2014). But they do suggest that they should consider that evidence more seriously. The approach also emphasizes evolutionary thinking as an aid for theory construction, a style of speculative theorizing available to philosophers.

### ACKNOWLEDGEMENTS

I submitted for discussion an early draft of this paper in the *Intentionality* research group and received significant feedback to continue writing. I presented a later version at the Workshop on Animal Minds, where I benefitted greatly from comments offered by Andrés Crelier and Laura Danón. The research for this paper was possible through a Ph.D. scholarship granted by CONICET (National Scientific and Technical Research Council). Finally, I want to thank my late aunt Clara Bauer, who taught me with patience and love what I know about the English language.

### REFERENCES

- Allen, Colin (2014). Models, Mechanisms and Animal Minds: Models, Mechanisms, and Animal Minds. *The Southern Journal of Philosophy*, 52, 75-97. <https://doi.org/10.1111/sjp.12072>
- Andrews, Kristin (2020). Naïve Normativity: The Social Foundation of Moral Cognition. *Journal of the American Philosophical Association*, 6(1), 36-56. <https://doi.org/10.1017/apa.2019.30>
- Bicchieri, Cristina (2016). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190622046.001.0001>
- Bird, Christopher David and Nathan John Emery (2009). Rooks Use Stones to Raise the Water Level to Reach a Floating Worm. *Current Biology*, 19(16), 1410-14. <https://doi.org/10.1016/j.cub.2009.07.033>
- Brosnan, Sarah F. and Frans B. M. de Waal (2003). Monkeys Reject Unequal Pay. *Nature*, 425(6955), 297-299. <https://doi.org/10.1038/nature01963>
- Buckner, Cameron (2013). Morgan's Canon, Meet Hume's Dictum: Avoiding Anthropofabulation in Cross-Species Comparisons. *Biology & Philosophy*, 28(5), 853-871. <https://doi.org/10.1007/s10539-013-9376-0>

- Cummins, Robert (2010). *The World in the Head*. Oxford/New York: Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199548033.001.0001>
- Currie, Adrian (2021). *Comparative Thinking in Biology*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108616683>
- Danón, Laura (2019). Animal Normativity. *Phenomenology and Mind*, (17), 176-187. <https://doi.org/10.13128/PAM-8035>
- Emery, Nathan J. and Nicola S. Clayton (2004). The Mentality of Crows: Convergent Evolution of Intelligence in Corvids and Apes. *Science*, 306(5703), 1903-1907. <https://doi.org/10.1126/science.1098410>
- Fitzpatrick, Simon (2017). Against Morgan's Canon. In Kristin Andrews and Jacob Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* (pp. 437-447). London: Routledge. <https://doi.org/10.4324/9781315742250-43>
- Fitzpatrick, Simon (2020). Chimpanzee Normativity: Evidence and Objections. *Biology & Philosophy*, 35(4), 45. <https://doi.org/10.1007/s10539-020-09763-1>
- Greenberg, Gary and Maury M. Haraway (1998). *Comparative Psychology: A Handbook*. London: Routledge. <https://doi.org/10.4324/9780203826492>
- Jelbert, Sarah A., Alex H. Taylor, Lucy G. Cheke, Nicola S. Clayton and Russell D. Gray (2014). Using the Aesop's Fable Paradigm to Investigate Causal Understanding of Water Displacement by New Caledonian Crows. *PLoS ONE*, 9(3), 1-9. <https://doi.org/10.1371/journal.pone.0092895>
- Kelly, Daniel and Stephen Setman (2021). The Psychology of Normative Cognition. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2021/entries/psychology-normative-cognition/>
- Luncz, Lydia V. and Christophe Boesch (2014). Tradition over Trend: Neighboring Chimpanzee Communities Maintain Differences in Cultural Behavior Despite Frequent Immigration of Adult Females. *American Journal of Primatology*, 76(7), 649-657. <https://doi.org/10.1002/ajp.22259>
- Luncz, Lydia V., Giulia Sirianni, Roger Mundry y Christophe Boesch (2018). Costly Culture: Differences in Nut-Cracking Efficiency between Wild Chimpanzee Groups. *Animal Behaviour*, 137, 63-73. <https://doi.org/10.1016/j.anbehav.2017.12.017>

- Monsó, Susana and Kristin Andrews (2022). Animal Moral Psychologies. In Manuel Vargas and John M. Doris (Eds.), *The Oxford Handbook of Moral Psychology* (pp. 388-420). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198871712.001.0001>
- Olmstead, Mary C. and Valerie A. Kuhlmeier (2022). Comparative Cognition. In Patrizia d’Ettorre, Amanda R. Ridley, and Todd M. Freeberg (Eds.), *The Routledge International Handbook of Comparative Psychology* (pp. 84-95). New York: Routledge. <https://doi.org/10.4324/9781003091868>
- Papini, Mauricio R. (2021). *Comparative Psychology: Evolution and Development of Brain and Behavior*. New York: Routledge. <https://doi.org/10.4324/9781003080701>
- Povinelli, Daniel J. (1993). Reconstructing the Evolution of Mind. *American Psychologist*, 48(5), 493-509. <https://doi.org/10.1037/0003-066X.48.5.493>
- Roth, Martin and Robert Cummins (2017). Neuroscience, Psychology, Reduction, and Functional Analysis. In David M Kaplan (Ed.), *Explanation and Integration in Mind and Brain Science* (pp. 29-43). New York: Oxford University Press. <https://doi.org/10.1093/oso/9780199685509.003.0002>
- Roughley, Neil and Kurt Bayertz (Eds.) (2019). *The Normative Animal? On the Anthropological Significance of Social, Moral, and Linguistic Norms*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190846466.001.0001>
- Schlingloff, Laura and Richard Moore (2017). Do Chimpanzees Conform to Social Norms? In Kristin Andrews and Jacob Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* (pp. 381-389). London: Routledge. <https://doi.org/10.4324/9781315742250-37>
- Schmidt, Marco F. H. and Hannes Rakoczy (2019). On the Uniqueness of Human Normative Attitudes. In Neil Roughley and Kurt Bayertz (Eds.), *The Normative Animal? On the Anthropological Significance of Social, Moral, and Linguistic Norms* (pp. 121-135). Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190846466.003.0006>
- Sherry, David F. (2005). Do Ideas about Function Help in the Study of Causation? *Animal Biology*, 55(4), 441-456. <https://doi.org/10.1163/157075605774840950>
- Shettleworth, Sara J. (2009). The Evolution of Comparative Cognition: Is the Snark Still a Boojum? *Behavioural Processes*, 80(3), 210-217. <https://doi.org/10.1016/j.beproc.2008.09.001>

- Sober, Elliot (2005). Comparative psychology meets evolutionary biology. In L. Daston and G. Mitman (Eds.), *Thinking with animals: New perspectives on anthropomorphism* (pp. 85-99). New York: Columbia University Press.
- Sterelny, Kim (2019). Norms and Their Evolution. In Tracy B. Henley, Matt J. Rossano, and Edward P. Kardas (Eds.), *Handbook of Cognitive Archaeology* (pp. 375-397). London: Routledge. <https://doi.org/10.4324/9780429488818-20>
- Sterelny, Kim (2021). *The Pleistocene Social Contract: Culture and Cooperation in Human Evolution*. New York: Oxford University Press. <https://doi.org/10.1093/oso/9780197531389.001.0001>
- Tinbergen, N. (1963). On Aims and Methods of Ethology. *Zeitschrift Für Tierpsychologie*, 20(4), 410-433. <https://doi.org/10.1111/j.1439-0310.1963.tb01161.x>
- Tomasello, Michael (2016). *A Natural History of Human Morality*. Cambridge (Mass.): Harvard University press. <https://doi.org/10.4159/9780674915855>
- Van Schaik, Carel P. (2012). Animal Culture: Chimpanzee Conformity? *Current Biology*, 22(10), R402-4. <https://doi.org/10.1016/j.cub.2012.04.001>
- Vincent, Sarah, Rebecca Ring and Kristin Andrews (2019). Normative Practices of Other Animals. In Aaron Zachary Zimmerman, Karen Jones, and Mark Timmons (Eds.), *The Routledge Handbook of Moral Epistemology* (pp. 57-83). New York: Routledge. <https://doi.org/10.4324/9781315719696-4>
- Vonk, Jennifer and Todd K. Shackelford (Eds.) (2012). *The Oxford Handbook of Comparative Evolutionary Psychology*. Oxford: Oxford University Press.
- Waal, Frans B. M. de (1989). *Peacemaking Among Primates*. Cambridge, Mass: Harvard University Press. <https://doi.org/10.4159/9780674033085>
- Waal, Frans B. M. de (1991). Complementary Methods and Convergent Evidence in the Study of Primate Social Cognition. *Behaviour*, 118(3/4), 297-320. <https://www.jstor.org/stable/4534969>
- Waal, Frans B. M. de (2007). *Chimpanzee Politics: Power and Sex Among Apes*. Baltimore, Maryland: Johns Hopkins University Press.
- Waal, Frans B. M. de (2014). Natural Normativity: The 'Is' and 'Ought' of Animal Behavior. *Behaviour*, 151(2-3), 185-204. <https://doi.org/10.1163/1568539X-00003146>
- Waal, Frans B. M. de (2016). *Are We Smart Enough to Know How Smart Animals Are?* New York: W. W. Norton & Company.

- Waal, Frans B. M. de, Stephen Macedo, Josiah Ober and Robert Wright (2006). *Primates and Philosophers: How Morality Evolved*. The University Center for Human Values Series. Princeton, N.J: Princeton University Press.
- Wasserman, Edward A. and Thomas R. Zentall (2009). *Comparative Cognition: Experimental Explorations of Animal Intelligence*. New York: Oxford university press.
- Westra, Evan and Kristin Andrews (2022). A Pluralistic Framework for the Psychology of Norms. *Biology & Philosophy*, 37(5), 40. <https://doi.org/10.1007/s10539-022-09871-0>



# INVASIÓN SILENCIOSA: LA PRIMATOLOGÍA DE IMANISHI Y EL SESGO CULTURAL EN LA CIENCIA<sup>1</sup>

## TRADUCCIÓN

Frans B. M. DE WAAL  
*Living Links, Yerkes Primate Center, Emory University, United States*  
dewaal@emory.edu

### 1. INTRODUCCIÓN

Cuando se trata de nuestra relación con la naturaleza, no hay forma de escapar de la tensión entre percepción y proyección. A menudo, lo que descubrimos en la naturaleza es lo que antes pusimos en ella. En consecuencia, la forma en que los naturalistas han contribuido a la misión de 'la humanidad ha de conocerse a sí misma', sólo puede entenderse en el contexto del cristal con que se mire el espejo de la naturaleza. Dado que no nos es posible quitarnos los cristales de esas gafas, la segunda mejor opción que nos queda es comparar otras alternativas.

El presente ensayo explora los prejuicios culturales en el marco de mi pequeño rincón de la ciencia, el comportamiento de los monos y de los simios. Dado que la forma en que miramos a otros animales refleja la forma en que nos miramos a nosotros mismos, el estudio del

1. Traducción: E. Joaquín Suárez-Ruiz y Carolina Scotto. Esta traducción fue autorizada por su autor, Frans de Waal, y contempla las normas de copyright de su editorial original, *Animal Cognition*. © Springer-Verlag 2003.

comportamiento animal está sujeto a preconceptos culturales mucho mayores que los que podrían poseer otros campos de las ciencias naturales. Por ejemplo, podemos considerar a los organismos como empresas cooperativas, tanto en el interior (entre las células del cuerpo), como en el exterior (cuando los animales cooperan para sobrevivir). Pero también se puede hacer hincapié en la competencia despiadada y en los llamados genes 'egoístas'. No es difícil apoyar cualquiera de las dos posiciones. No obstante, en Occidente nos encanta representar la naturaleza con garras y dientes.

El fundador de la primatología japonesa, Kinji Imanishi (1902-1992), concebía a la naturaleza como inherentemente armoniosa. Las especies encajaban en un gran todo orgánico y cada especie encontraba su propio nicho. Tijs Goldschmidt, un biólogo holandés que trabaja con cíclidos en el lago Victoria, me contó que el contraste con el enfoque occidental persiste aún hoy (Goldschmidt, 1998). Goldschmidt visitó una vez a colegas japoneses que estudiaban la misma familia de peces que él, en el cercano lago Tanganica. Mientras que el equipo de Goldschmidt explicó la explosión de especies en estos lagos en términos de competencia y explotación mutua, los japoneses la comprendieron en términos de roles complementarios a nivel ecosistémico. Si bien ambos equipos coincidieron a nivel de los datos, operaron sobre la base de perspectivas sorprendentemente diferentes.

Los desacuerdos entre Oriente y Occidente respecto de la competencia y la cooperación como aspectos de la naturaleza, se remontan al menos al debate de finales del siglo XIX entre Thomas Henry Huxley y Piotr Kropotkin. Mientras que el primero adoptó una visión "gladiadora" de la naturaleza, el segundo defendió un modelo más sinérgico (Todes, 1989; de Waal, 1996). Estos desacuerdos rara vez dan lugar a un claro ganador. Más bien, tienden a tener un tono de debate en clave de 'el vaso medio lleno' versus 'el vaso medio vacío'.

Un ejemplo destacado de tal discusión fue la reacción a las opiniones de Imanishi por parte de la paleontóloga británica Beverly Halstead, quien consideró necesario viajar hasta Japón para dejarle las cosas claras al viejo maestro. En 1984, armada con un montón de prejuicios y sin admitir de primera mano ningún conocimiento de los escritos de Imanishi, Halstead fue a enfrentarlo. En su manuscrito inédito en inglés (cuya copia se puede encontrar en la Biblioteca de la Universidad de Kyoto), señaló: "A mi manera occidental, vine a Kyoto, el hogar de Imanishi y su escuela, buscando al hombre y sus ideas, pero he venido como una oponente declarada" (Halstead, 1985, manuscrito inédito). Después de entregarle un regalo al profesor emérito de 82 años (una botella de whisky), le entregó



un documento traducido al japonés que contenía afirmaciones como “La teoría de la evolución de Imanishi es japonesa en su irrealidad” y “Usted ve la madera, pero los árboles quedaron fuera de foco”. No es de extrañar que el rostro de Imanishi, como recordó Halstead, revelara un profundo arrepentimiento de haber aceptado la reunión ¿Qué pudo haber impulsado a Halstead a ser tan grosera?

¿Por qué, luego de su regreso a Gran Bretaña, escribió un artículo que destrozaba no sólo las opiniones de Imanishi, sino también las de todo su país? ¿Cómo se atrevió *Nature* a publicarlo con la condescendiente introducción que sigue: “La popularidad de los escritos de Kinji Imanishi en Japón ofrece una visión interesante de la sociedad japonesa” (Halstead, 1985)? ¿No podría aplicarse el mismo subtítulo a, digamos, la teoría de Darwin? Como se ha señalado ya muchas veces, no puede ser una coincidencia que las ideas sobre el capitalismo de libre mercado y la lucha por la existencia surgieran al mismo tiempo y en el mismo lugar. El hábito común de formular las cuestiones evolutivas en términos de costos y beneficios, deja pocas dudas sobre esta conexión.

Lo que tenemos aquí, entonces, es el caso familiar de una cultura que percibe los sesgos de la otra de un modo más crítico que como percibe los suyos propios. Incluso si hoy en día las ideas ecológicas y evolutivas de Imanishi se considerasen problemáticas, él y sus seguidores tenían razón en bastantes puntos. Esto se desprende, con particular claridad, del cambio fundamental ocurrido durante últimas décadas en el estudio del comportamiento animal: los científicos occidentales, aunque a menudo ignoren su origen, han ido adoptando, en gran escala, conceptos y enfoques japoneses. No está claro si también han adoptado la perspectiva subyacente, que está bastante alejada de la suya propia, pero ciertamente han sido receptivos a sus técnicas de observación y a sus conceptos específicos, como las relaciones individualizadas y la transmisión cultural, empleados por primera vez en la primatología japonesa.

## 2. EL LEGADO DE IMANISHI

Imanishi fue un autor extraordinariamente prolífico y ampliamente reconocido en el campo de las ciencias biológicas: el Stephen Jay Gould de Japón. Comenzó como entomólogo, pero también fue ecólogo, antropólogo, primatólogo, montañista y filósofo. Recibió un puesto oficial en la facultad (en humanidades, no en ciencias), recién cuando tenía aproximadamente 50 años. Al ser de ascendencia acaudalada, podía hacer lo que quisiera sin las obligaciones que conlleva mantener un salario.

Tan sólo tenía una habitación en la Universidad de Kyoto sin más muebles que un escritorio bajo en el que escribía sus libros, sentado en posición de loto sobre un tatami. Era un hombre ascético, culto y de inmensa influencia.

Además de escalar el Himalaya, Imanishi tenía dos intereses principales. Uno era la interconexión entre todos los seres vivos y los ambientes en los que se encuentran. Aunque rara vez mencionó a quienes influyeron en él, era una persona muy leída y elementos de su enfoque se pueden rastrear en influencias externas, que van desde Jacob von Uexküll hasta Petr Kropotkin y, quizás, sobre todo, Kitaro Nishida, el fundador de una escuela filosófica que fue particularmente influyente en las décadas de 1930 y 1940. No puedo juzgar esto por mí mismo, ya que mi información proviene de fuentes secundarias, pero el énfasis de Imanishi en la intuición y la percepción del todo, su aversión al reduccionismo y su visión de que el individuo es secundario respecto de la sociedad, probablemente derivaron de Nishida, el filósofo de la 'nada' de Kioto, quien solía tener profundos pensamientos mientras paseaba por un pequeño río rústico bordeado de cerezos –todavía conocido como el camino del filósofo–, que pasa por el campus universitario (Yoshimi, 1998)<sup>2</sup>.

Imanishi realizó la mayor parte de su investigación sobre larvas de efímera en el río Kamogawa, mucho más grande que el anterior, y que atraviesa el corazón de Kioto. Su trabajo sobre la vida acuática lo llevó a desarrollar la noción de 'segregación de hábitat', que refiere a que especies diferentes pero relacionadas seleccionan sus propios estilos de vida y microhábitats, lo que les permite coexistir armoniosamente en el mismo entorno. Imanishi no intentó explicar cómo pudo haberse producido la segregación y se opuso vehementemente a las explicaciones que implicaran conflictos.

El segundo interés y legado duradero de Imanishi se relaciona con el estudio del comportamiento de los primates. En este caso, su enfoque fue particularmente innovador gracias a la ausencia del dualismo

2. En el otoño de 1998 visité China y Japón durante varios meses con una beca otorgada por la Sociedad Japonesa para la Promoción de la Ciencia (JSPS, por sus siglas en inglés). Durante este tiempo, hablé con muchos primatólogos y otros académicos tanto en mi base de operaciones, Kioto, como en todo Japón, desde el norte hasta el sur. Esto incluyó una conversación con el fallecido Jun'ichiro Itani, quien me ofreció una visión de primera mano de las ideas tempranas de Imanishi y su recepción tanto en Japón como en Occidente. También incluyó una visita a Koshima y conversaciones con Satsue Mito, quien desde el principio ayudó en los estudios sobre el lavado de patatas en la isla. Se remite al lector a De Waal (2001), para una descripción más completa de mis impresiones sobre la ciencia japonesa y las conexiones y diferencias con el enfoque occidental.

humano-animal. Siendo producto de una cultura que no presenta a la especie humana como la única con alma, Imanishi no tuvo problemas ni con la idea de la evolución ni con la de los humanos como descendientes de los simios. Para la mente budista, esto no sólo resulta eminentemente plausible, sino incluso probable y no tiene nada de insultante (Asquith, 1991; Sakura, 1998; Matsuzawa, 2003).

La 'gran cadena del ser' de Platón, que coloca a los humanos por encima de todos los demás animales, está ausente en la filosofía oriental. En la mayoría de los sistemas de creencias orientales, el alma humana puede reencarnar en muchas formas, por lo que todos los seres vivos están espiritualmente conectados. Un hombre puede convertirse en un pez y un pez puede convertirse en un dios o una diosa. El hecho de que los primates, nuestros parientes animales más cercanos, sean nativos de China y Japón, sólo ha ayudado a fortalecer la creencia en la interconexión de la vida. A diferencia de las fábulas europeas, que están pobladas de cuervos, conejos, zorros y similares, los cuentos populares y la poesía japoneses y chinos están salpicados de referencias a gibones y monos. Ejemplo de ellos son los tres macacos sabios del budismo Tendai ("no veas el mal, no escuches el mal, no hables el mal").

Sentir humildad hacia los animales afecta la forma en que los estudiamos. El estudio del comportamiento animal en Japón nunca ha estado contaminado por sentimientos de superioridad o de aversión a reconocer características humanas en ellos. Según el estudiante más respetado de Imanishi, el fallecido Junichiro Itani (1926-2001): "Dado que la cultura japonesa no enfatiza la diferencia entre personas y animales, está relativamente libre del hechizo del antiantropomorfismo... Creemos que esto ha llevado a muchos descubrimientos importantes" (Itani, 1985; ver también Asquith, 1984, 1986). Así, los primatólogos japoneses trazaron las relaciones de parentesco a lo largo de múltiples generaciones, asumiendo que los animales poseen una vida familiar compleja, tal como la tenemos nosotros. Comenzaron con ello incluso antes de que los científicos occidentales siquiera lo pensarán (por ejemplo, Kawamura, 1958; Yamada, 1963), y años antes de que William Hamilton (1964) desarrollara la teoría de la 'selección de parentesco'. Las redes de parentesco fueron un verdadero descubrimiento, quizás el mayor de la primatología japonesa (Reynolds, 1992).

De hecho, la aceptación sin dificultades de un aspecto importante de la teoría de la evolución –la continuidad entre todas las formas de vida– significó que, desde un principio, las cuestiones sobre el comportamiento animal no estuvieran contaminadas por la división entre

humanos y animales asumida en Occidente. Como resultado, los estudiantes de Imanishi pudieron avanzar rápidamente con una agenda claramente antropológica: al estudiar a otros primates, intentaron comprender los orígenes de la familia y la sociedad humanas. En todo esto, Imanishi estaba muy por delante del célebre paleontólogo occidental Louis Leakey, quien desarrolló una agenda similar. Leakey envió a varios primatólogos a estudiar los grandes simios en libertad, con la creencia de que estos animales podrían proporcionarnos información sobre las primeras etapas de la evolución humana. No obstante, cuando lo hizo, en las décadas de 1960 y 1970, las preguntas y técnicas que resultarían útiles en tal esfuerzo ya habían sido desarrolladas por primatólogos japoneses, quienes habían identificado individualmente a sus monos y los habían seguido el tiempo suficiente para comprender la extraordinaria complejidad de la sociedad primate, así como también el grado de diferencia entre cada grupo. Lo más importante es que Imanishi había formulado la cuestión de la cultura animal de una manera que invitaba a realizar más estudios (véase la siguiente discusión). Ahora bien, en lugar de comparar a Imanishi con Leakey, el paralelo más apropiado es con Ray Carpenter, el primatólogo estadounidense pionero. Carpenter era un fisiólogo de formación, pero también un científico del comportamiento de primer nivel que prefería el campo al laboratorio. Trabajó con *macacos rhesus* liberados en la isla caribeña de Cayo Santiago, así como también con monos aulladores y gibones salvajes. Estaba interesado en las relaciones sociales y dibujó sociogramas que trazaban la estructura del grupo (Carpenter, 1964). Si bien no llegó tan lejos en esto como los primatólogos japoneses, quienes pudieron distinguir más de 100 monos y rastrear sus vínculos familiares a lo largo de generaciones, Carpenter compartía con ellos una perspectiva claramente “sociológica”. No sorprende, por tanto, que Carpenter fuera el primer occidental en convertirse en un firme partidario de la primatología japonesa (véase la figura 1).

El enfoque de Imanishi sobre el comportamiento de los primates equivale a un cambio paradigmático que hoy ha sido adoptado por toda la primatología y más allá. Por lo tanto, si ya no percibimos el antropomorfismo como el problema que alguna vez fue (Mitchell *et al.*, 1997; de Waal, 1999), y si los estudiantes de animales longevos –ya sea que observen delfines o elefantes– identifican de manera rutinaria individuos y los siguen a lo largo de su vida (de Waal y Tyack, 2003), es porque estamos empleando técnicas de Oriente que inicialmente fueron objeto de burla y resistidas por Occidente.



**Figura 1.** Kinji Imanishi (1902-1992) y Ray Carpenter (1905-1975), frente a la Universidad Estatal de Pensilvania, donde Carpenter estuvo trabajando durante la visita de Imanishi a Estados Unidos. Carpenter presenta el primer número de la revista *Primates*, la más antigua en su campo, ahora publicada por Springer-Verlag. La fotografía fue tomada por el fallecido Itani en 1958. Copyright: Archivos visuales en memoria de Jun'ichirou Itani del Instituto de Investigación de Primates de la Universidad de Kioto. Con permiso especial de Tetsuro Matsuzawa

### 3. LA BARRERA DEL IDIOMA

En 1958, Imanishi y sus estudiantes recorrieron universidades estadounidenses para informar de sus hallazgos. Siendo que humanizaban a sus sujetos, se encontraron con muchas burlas y con un profundo escepticismo respecto de la capacidad de simples humanos a la hora de distinguir a todos esos monos. A la gente le costaba creer que tal hazaña fuera posible y expresaban incredulidad frente a sus visitantes (Itani, comunicación personal). Tampoco debemos olvidar que hace tan sólo unas décadas los profesores universitarios occidentales advertían a los estudiantes de primatología contra el enfoque ateórico, el antropomorfismo y la falta general de relevancia de los artículos de sus colegas japoneses, mientras desalentaban las referencias a esta literatura (Asquith, 1996).

Entonces, ¿cómo es posible que los principios básicos de la escuela de Imanishi se den ahora por sentados en Occidente? Para comprender cómo esta "invasión alienígena" de ideas pudo haber tenido lugar frente a nuestras narices, debemos mirar la cultura oriental y también apreciar cómo el monopolio lingüístico afecta a la ciencia (Gibbs, 1995).

La respuesta a la primera pregunta es, como hemos visto, que la ciencia oriental no tenía ninguna afición por el tradicional dualismo humano/animal occidental. Las ventajas de ignorar este dualismo resultaron inmediatamente obvias para los científicos de mente abierta, como Carpenter, quienes ayudaron a acelerar un proceso en Occidente que podría haber ocurrido de todos modos.

La respuesta a la segunda pregunta está en el lenguaje (Bartholomew, 1998; Asquith, 2000). Para quienes no hablan inglés, resulta difícil hacerse oír en un mundo angloparlante. Dado que el inglés no es mi lengua materna, estoy familiarizado con el esfuerzo que implica escribir y hablar en otro idioma, aunque mi holandés nativo es probablemente el idioma más cercano al inglés. Cuando era estudiante, todos mis libros de texto estaban en dos idiomas extranjeros (alemán e inglés), y cuando más tarde comencé a escribir artículos, pasé una cantidad excesiva de tiempo buscando términos en gruesos diccionarios que permitieran expresar mis pensamientos en inglés. Este esfuerzo, que para los científicos japoneses seguramente se multiplique por diez, es despreocupadamente ignorado por los hablantes nativos de inglés.

El típico hablante nativo de inglés es monolingüe. La falta de familiaridad con otros idiomas le hace imaginar que deben ser copias del inglés. No obstante, idiomas distintos no sólo tienen diferentes palabras y gramáticas, sino que también representan diferentes visiones del mundo. Al ser conceptualmente diferentes, muchas expresiones y matices resultan intraducibles. Por ejemplo, el holandés está plagado de diminutivos (inexistentes en la mayoría de los idiomas) que transmiten una ternura propia de un país pequeño y ordenado. La forma en que los franceses hablan de comida, por poner otro ejemplo, ni siquiera es imaginable para la mayoría de los angloparlantes. Las diferencias, sin embargo, son mucho más profundas: determinan la forma en que construimos la realidad. Por lo tanto, si las palabras chinas a menudo sirven a la vez como sustantivos y como verbos, esto hace que sea natural para los chinos ver los objetos como eventos y entender el mundo como si estuviera formado por procesos en lugar de por entidades.

El inglés, que por un accidente de la historia se ha convertido en la lengua dominante del mundo, es un idioma que está perfectamente bien. Tener un solo idioma para ponencias y conferencias internacionales, en

mi opinión, es preferible a tener varios idiomas en competencia. Entonces, el inglés en sí no es el problema: el problema es la actitud de los angloparlantes nativos.

Naturalmente, hablas tu propio idioma más rápido que cualquier otro. Esto puede hacer imposible que quienes no son hablantes nativos del inglés puedan seguir fluidamente las discusiones en los encuentros internacionales. Esta situación empeora en aquellas ocasiones en las que un angloparlante no se anda con rodeos cuando debate con un científico cuyo inglés es pobre. He visto que eso sucede a menudo. El angloparlante se levanta del público, articula una pregunta punzante, a veces mezclada con un chiste, y luego apenas se toma el tiempo de escuchar la respuesta torpemente formulada de su oponente. Al ser los angloparlantes nativos los que dominan las discusiones en general, se presentan como grandes mentes pavoneándose con la seguridad de que nadie se atreverá a desafiarlos.

Las buenas ideas formuladas en un mal inglés mueren o son reempaquetadas. Si esto último ocurre, una vez que la idea ha pasado al dominio del buen inglés, su origen tiende a olvidarse. Es un poco como la interpretación cinematográfica de una obra de teatro francesa (por ejemplo, “La jaula de pájaros” – *La Cage aux Folles*; “La Bella y la Bestia” – *La Belle et la Bête*): una vez que se ha estrenado la película, la gran mayoría de la gente cree que la idea surgió en Hollywood. Una vez expresada en inglés, la idea en cuestión se vuelve inglesa o americana. Este es un proceso natural que probablemente se aplique a cualquier idioma, pero en ciencia deberíamos dar crédito intelectual a quien se lo merece.

Entonces, una de las razones por las que el pensamiento oriental pudo infiltrarse en el estudio del comportamiento animal sin ser advertido reside en que se filtró en la literatura a través de formulaciones y traducciones enrevesadas que los angloparlantes nativos encontraron fácil de mejorar. En el proceso procedieron a borrar parte del crédito de las nuevas ideas. Por lo tanto, aunque Imanishi nos haya encaminado en el sendero de la cultura animal –que ahora es el tema más candente en nuestro campo– los científicos occidentales rara vez, o nunca, mencionan su nombre en este contexto.

#### 4. CULTURA ANIMAL

Ya en 1952, cuando los etólogos europeos trabajaban en teorías del instinto y los conductistas estadounidenses recompensaban ratas por presionar palancas, Imanishi escribió un artículo en el que criticaba

las opiniones establecidas sobre los animales (Imanishi, 1941, 1952). Incluyó un debate entre una avispa, un mono, un evolucionista y un lego, en el que se planteó la posibilidad de que, además de nosotros, otros animales también pudieran tener cultura. Hirata *et al.* (2001) proporcionan una traducción de una parte de este debate imaginario. La definición propuesta era simple: si los individuos aprenden unos de otros, su comportamiento puede, con el tiempo, llegar a ser diferente del de otros grupos, creando así una cultura característica (Itani y Nishimura, 1973; Nishida, 1987).

Este enfoque redujo la cultura a su mínimo común denominador: la transmisión social, más que a la genética del comportamiento. Esto se confirmó pocos años después de la publicación del libro, a partir de observaciones de macacos japoneses lavando batatas en la isla de Koshima. Ahora sabemos que el aprendizaje cultural está muy extendido e incluye el canto de los pájaros (p. ej., West *et al.*, 2003), el uso de herramientas por parte de los chimpancés (p. ej., Whiten *et al.* 1999) y las técnicas de caza de las ballenas (p. ej., Rendell y Whitehead, 2001). Nuevos ejemplos se descubren casi a diario.

Hasta hace poco, sin embargo, los científicos occidentales se habían resistido a la idea de la cultura animal, principalmente insistiendo en mecanismos muy específicos de transmisión social –como la enseñanza y la imitación– que muchos animales pueden no manifestar, y de los cuales incluso la cultura humana puede no depender en el grado que estos autores suponen (Premack y Premack, 1994; Tomasello, 1994). En un desafío directo a la escuela de Imanishi, Galef (1992) escribió una influyente crítica a los estudios de Koshima en la que afirmaba que (1) el aprovisionamiento de alimentos en la isla puede haber sido realizado selectivamente para recompensar a los individuos que mostraban el comportamiento deseado, como lavar papas, y (2) el aprendizaje individual podría ser suficiente como explicación de la propagación del hábito en cuestión. Galef, que nunca puso un pie en Koshima, se basó en un par de frases de Green (1975), quien sí había visitado la isla en 1968 y en 1969.

En ese momento, el personal de Koshima ocasionalmente admitía a personas que querían ir a ver a los monos lavando papas. Para ello, alimentaban a los monos cerca del océano y se aseguraban de que los mejores “ejecutantes” estuvieran disponibles. Green (1975), que asistió sólo a una de esas escenas de recolección de papas durante toda su estancia, estaba consciente de que este método de aprovisionamiento pretendía beneficiar a los turistas y a los investigadores visitantes, incluido él mismo. Su relato anecdótico de más de 15 años después del inicio del lavado de papas no puede decirnos cómo se originó o se extendió



tal hábito. Para ello, deberíamos recurrir a la cuidadosa documentación de Kawai (1965), que cubrió un período bastante anterior durante el cual pocos forasteros se acercaron a Koshima.

Mis propias conversaciones con Satsue Mito (quien dirigió el suministro de papas en los primeros años), me generaron serias dudas de que los procedimientos imaginados por Galef (1992) hubiesen sido aplicados alguna vez (de Waal, 2001). En primer lugar, no habría sido lógico. Para que un comportamiento se propague, es de vital importancia que los individuos que no exhiben el comportamiento en cuestión tengan la oportunidad de hacerlo. De ahí, justamente, la necesidad de proporcionar papas a los monos no lavadores. En segundo lugar, uno no podría alimentar a una manada de monos como a uno se le ocurra sin causar, al mismo tiempo, una increíble confusión. Para que ello no ocurra es necesario “alimentar” la jerarquía. Si uno alimentara a los monos jóvenes y de bajo rango antes que al resto, los superiores los acosarían hasta el punto que pondrían en peligro su vida. Mito, habiendo entendido esto, alimentó primero a los machos adultos y a las hembras de alto rango. No obstante, a pesar de esta técnica de alimentación, el lavado de papas comenzó con los individuos jóvenes y de bajo rango. De hecho, aunque los machos de mayor edad fueron los primeros en ser alimentados, nunca aprendieron a lavar. A su vez, a partir de las excelentes revisiones de Kawai (1965), Watanabe (1994), Hirata *et al.* (2001) y otros, resulta evidente que el lavado de patatas se extendió entre los monos de una forma que era consistente con las relaciones sociales de la manada. Por lo tanto, los primeros individuos en manifestar el comportamiento después de que una hembra joven, Imo, comenzara con el hábito, fueron sus coetáneos y su madre ¿Cuál sería la probabilidad de que esto sucediese por casualidad? En un análisis *post hoc* de los datos existentes, la celeridad del aprendizaje parecía consistente con un modelo de transmisión social. Lefebvre (1995) argumentó que el aprendizaje social puede distinguirse del aprendizaje individual por la rapidez con la que se propaga un nuevo comportamiento. Es posible que el aprendizaje social produzca una función aceleradora, ya que, mientras que el aumento en el número de practicantes del nuevo comportamiento aumenta la probabilidad de que los individuos que no lo habían practicado antes queden expuestos a él, se espera que el aprendizaje individual produzca una función lineal. Basándose en estos supuestos, Lefebvre concluyó que las tradiciones primates más analizables eran consistentes con el aprendizaje social. Por lo tanto, no hay ninguna razón urgente para renunciar al lavado de papas en Koshima como ejemplo de transmisión social de un hábito específico en una población de monos (de Waal, 2001).

Esto no conlleva afirmar que la cuestión planteada por Galef *et al.* sobre el mecanismo de transmisión sea irrelevante. Mientras que los biólogos consideran que el mecanismo es secundario a la función, ciertamente en el nivel de la definición de un fenómeno (por ejemplo, no definimos la locomoción o la respiración por la manera en que se logra), las diversas formas en que la transmisión cultural se efectiviza siguen siendo enigmáticas y exigen total atención. Sin embargo, hay que añadir que los psicólogos del aprendizaje, quienes provienen de una tradición en la que se testean animales individuales en tareas artificiales, tal vez no posean el mejor enfoque disponible para abordar el enigma mencionado. La confiada afirmación de Galef (1992) de que los actos imitados se extinguirán automáticamente si no consiguen generar recompensas, permite ilustrar dicha limitación. Si bien esto resulta lógico desde la perspectiva del refuerzo, conocemos muchos ejemplos de conductas culturalmente transmitidas que no poseen ningún valor de recompensa. Dos ejemplos son el manejo de piedras por parte de los monos japoneses en Arashiyama, que parece ser una actividad inútil (Huffman, 1996), y los intentos de cascar nueces por parte de chimpancés jóvenes. Respecto de este último caso, si bien los chimpancés bebés y jóvenes carecen de la habilidad, la fuerza y la coordinación para romper nueces, en sus juegos combinan piedras y nueces de una manera que eventualmente conduce a obtener el fruto comestible. Sin embargo, sus primeros aciertos se producen sólo después de muchos años de actividad sin recompensa alguna (Inoue-Nakamura y Matsuzawa, 1997). De hecho, contrariamente a lo que afirma Galef, el refuerzo extrínseco posee poca relevancia para la difusión de hábitos culturales (de Waal, 2001).

Un enfoque ecológicamente válido del aprendizaje social requeriría que consideráramos el contexto natural en el que tiene lugar. Por ejemplo, los experimentos de imitación que son populares actualmente, en los que primates no humanos u otros animales observan un modelo de otra especie (siempre la nuestra) o presencian una única demostración de un acto novedoso, están muy alejados de las circunstancias en las cuales evolucionó el aprendizaje social. En la naturaleza, los observadores y los modelos pertenecen a la misma especie y normalmente copian el comportamiento después de haberlo presenciado varias veces. Este es el tipo de contexto de aprendizaje sobre el que necesitamos saber más.

En lugar de centrarse en el refuerzo, las nuevas conceptualizaciones suponen que el aprendizaje social se basa en la identificación con modelos y en la necesidad de actuar como ellos. Tales supuestos conformistas subyacen al Aprendizaje Observacional basado en el Vínculo y

la Identificación (BIOL, por sus siglas en inglés) de De Waal (1998, 2001), así como también a la educación maestro-aprendiz de Matsuzawa *et al.* (2001). Estas teorías no tienen el problema de explicar por qué los experimentos de imitación con modelos humanos a veces fracasan: ellas asignan un papel crítico a la cercanía entre el observador y su modelo de comportamiento que, junto con el deseo de comportarse como los demás, promueve la convergencia conductual gracias a la cual se produce la cultura. Los mecanismos propuestos son tanto socioemocionales como cognitivos.

## 5. SUPUESTOS VERSUS TEORÍA

En sus esclarecedoras comparaciones entre la primatología occidental y la japonesa, Asquith señaló muchas distinciones útiles, como la ya mencionada falta de dualismo humano-animal en la religión oriental. Asquith no cree, como sí lo hizo Halstead (1985), que el darwinismo haya sido alguna vez rechazado por los científicos japoneses, sino más bien que, por un lado, los científicos occidentales tendieron a pensar en términos absolutos (es decir, o uno está a favor o está en contra de una idea) y, por otro lado, sus colegas japoneses simplemente se basaron en los elementos más atractivos de la teoría de la evolución, ignorando el resto (Asquith, 1986, 1991).

La primatología japonesa nunca fue tan ateórica como creían algunos científicos occidentales. Mi impresión es, más bien, que no todo el mundo se siente cómodo formulando ideas según la manera que es popular en Occidente. Establecemos teorías que son testeables, es decir, que pueden ser refutadas. Para nosotros, la ciencia es un proceso de confrontación que procura determinar quién tiene razón y quién no. Se necesita cierta mentalidad para operar de esta manera. Si bien a los científicos japoneses no les faltan suposiciones o expectativas sobre el mundo, son reacios a proponerlas de una manera que invite al desacuerdo.

Los primatólogos japoneses poseen expectativas claras, como se desprende de un ejemplo histórico poco conocido de la primatología de campo. Hasta finales de la década de 1960, la visión occidental era positivamente rousseauniana: los simios eran “nobles salvajes” autónomos, libres de lazos y obligaciones sociales. Lo único que hacían era desplazarse aleatoriamente de un árbol frutal a otro. Los siempre cambiantes grupos de chimpancés que los investigadores encontraban en los bosques de África parecían confirmar que carecían de una vida grupal

coherente. Mientras Jane Goodall describía a las chimpancés hembras y a sus crías dependientes como las únicas unidades vinculares (van Lawick-Goodall, 1968), un equipo japonés que trabajaba a sólo 130 kilómetros al sur de Gombe trabajaba bajo suposiciones bastante diferentes. Se preguntaron cómo sería posible que una especie que supuestamente llena el vacío entre nosotros y otros animales, carezca de una vida social compleja. Finalmente, a través de persistentes observaciones de campo, resolvieron el enigma y demostraron que los chimpancés viven en los denominados 'grupos unitarios' con una membresía estable (Nishida, 1968; Takasaki, 2000).

Los científicos occidentales pronto reemplazaron el término 'grupo unitario' por 'comunidad', lo que no facilitó la conservación del origen del descubrimiento. La sociedad unida por los machos [*male-bonded society*] del chimpancé es ahora un elemento básico del conocimiento primatólogo –existe amplia evidencia de agresión territorial entre diferentes comunidades (por ejemplo, Goodall, 1986)–, pero el descubrimiento inicial provino de una firme convicción de que los chimpancés no podían ser ni de lejos tan "individualistas" como la ciencia occidental los había presentado.

El punto importante de todo esto es que las opiniones de Imanishi, aunque no están formuladas como una teoría formal y han llegado a Occidente sólo con mucho retraso, claramente han prevalecido. Puntos de vista que entraban en conflicto con los dualismos occidentales tradicionales (por ejemplo, animal/humano, naturaleza/cultura) pasaron desapercibidos por nuestro pensamiento junto con técnicas de observación de gran valor. Esta invasión silenciosa desde Oriente ayudó a Occidente a deshacerse de parte de su bagaje cultural. Sin embargo, la forma como esto sucedió y la falta general de reconocimiento dan pistas de las dificultades que experimentan otros grupos culturales y lingüísticos cuando intentan encontrar una voz en la ciencia.

No debemos dejarnos engañar por la hegemonía científica de Occidente: es un pensamiento tan irreal como el de que un país pueda ordenar al resto del mundo. El estudio de la naturaleza no puede dejarse en manos de un solo sacerdocio en el que todos piensan igual. Cada cultura está demasiado absorta en su propia relación con la naturaleza como para poder dar un paso atrás y verla tal como es. Con el fin de obtener una imagen completa, se necesitan científicos con todo tipo de experiencia y que juntos asuman una tarea equivalente a la de comparar una variedad de imágenes en el laberinto de espejos de un parque de diversiones. En algún lugar de esa información tan distorsionada reside la verdad.

## AGRADECIMIENTOS

Este ensayo no podría haber sido escrito sin mis colegas japoneses durante mi agradable estancia en su país. Estoy inmensamente agradecido a Toshisada Nishida y Tetsuro Matsuzawa, quienes me invitaron a Kioto en 1998 y 2002, respectivamente, y al fallecido Junichiro Itani, Pamela Asquith, Toshikazu Hasegawa, Mariko Hiraiwa-Hasegawa, Michael Huffman, Suehisa Kuroda, Satsue Mito, Osamu Sakura, Hideko Takeshita, Kunio Watanabe, Soshichi Uchii y muchos otros colegas por sus útiles debates. También agradezco a los tres árbitros anónimos de esta revista por sus comentarios constructivos y referencias bibliográficas complementarias.

## REFERENCIAS

- Asquith, P. J. (1984). The inevitability and utility of anthropomorphism in description of primate behaviour. En R. Harré y V. Reynolds (Eds.), *The meaning of primate signals* (pp. 138-176). Cambridge: Cambridge University Press.
- Asquith, P. J. (1986) Anthropomorphism and the Japanese and Western traditions in primatology. En J. G. Else y P. C. Lee (Eds.), *Primate ontogeny, cognition, and social behavior* (pp. 61-71). Cambridge: Cambridge University Press.
- Asquith, P. J. (1991). Primate research groups in Japan: Orientations and East-West differences. En L. Fedigan & P. J. Asquith (Eds.), *The monkeys of Arashiyama. Thirty-five years of research in Japan and the West* (pp. 81-98). Albany, NY: SUNY Press.
- Asquith, P. J. (1996). Japanese science and western hegemonies. Primatology and the limits set to questions. En L. Nader (Ed.), *Naked science. Anthropological inquiry into boundaries, power, and knowledge* (pp. 239-256). New York: Routledge.
- Asquith, P. J. (2000). Negotiating science: internationalization and Japanese primatology. En S. Strum y L. M. Fedigan (Eds.), *Primate encounters* (pp. 165-183). Chicago: University of Chicago Press.
- Bartholomew, J. R. (1998). Japanese Nobel candidates in the first half of the twentieth century. *Osiris*, 13, 238-284. <https://doi.org/10.1086/649287>
- Carpenter, C. R. (1964). Characteristics of social behavior in nonhuman primates. En C. R. Carpenter (Ed.), *Naturalistic behavior of nonhuman primates* (pp. 358-364). University Park, Pennsylvania: Pennsylvania State University Press.

- Galef, B. G. (1992). The question of animal culture. *Hum. Nat.*, 3, 157-178. <https://doi.org/10.1007/BF02692251>
- Gibbs, W. W. (1995). Lost science in the Third World. *Sci. Am.*, 273, 92-99. <https://doi.org/10.1038/scientificamerican0895-92>
- Goldschmidt, T. (1998). *Darwin's dreampond*. Cambridge, Mass.: MIT Press.
- Goodall, J. (1986). *The chimpanzees of Gombe: patterns of behavior*. Cambridge, Mass.: The Belknap Press of Harvard University Press.
- Green, S. (1975). Dialects in Japanese monkeys: vocal learning and cultural transmission of locale-specific vocal behavior? *Z Tier-psychol*, 38, 304-314. <https://doi.org/10.1111/j.1439-0310.1975.tb02006.x>
- Halstead, L. B. (1985). Anti-Darwinian theory in Japan. *Nature*, 317, 587-589. <https://doi.org/10.1038/317587a0>
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. *J. Theor. Biol.*, 7, 1-52. [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4)
- Hirata, S., Watanabe, K. y Kawai, M. (2001). Sweet-potato washing revisited. En T. Matsuzawa (Ed.), *Primate origins of human cognition and behavior* (pp. 487-508). Tokyo/Heidelberg/Berlin: Springer. [https://doi.org/10.1007/978-4-431-09423-4\\_24](https://doi.org/10.1007/978-4-431-09423-4_24)
- Huffman, M. A. (1996). Acquisition of innovative cultural behaviors in non-human primates: a case study of stone handling, a socially transmitted behavior in Japanese macaques. En C. M. Heyes y B. G. Galef (Eds.), *Social learning in animals: the roots of culture* (pp. 267-289). San Diego: Academic. <https://doi.org/10.1016/B978-012273965-1/50014-5>
- Imanishi, K. (1941). *Japanese view of nature. The world of living things*. New York: Routledge Curzon.
- Imanishi, K. (1952). The evolution of human nature. En K. Imanishi (Ed.), *Ningen* (pp. 36-94). Tokyo: Mainichi-shinbunsha.
- Inoue-Nakamura, N. y Matsuzawa, T. (1997). Development of stone tool use by wild chimpanzees. *J. Comp. Psychol.*, 111, 159-173. <https://doi.org/10.1037/0735-7036.111.2.159>
- Itani, J. (1985). The evolution of primate social structures. *Man*, 20, 593-611. <https://doi.org/10.2307/2802752>
- Itani, J. y Nishimura, A. (1973). The study of infrahuman culture in Japan: a review. En E. W. Menzel (Ed.), *Precultural primate behavior* (pp. 26-50). Basel: Karger.
- Kawai, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima islet. *Primates*, 6, 1-30. <https://doi.org/10.1007/BF01794457>
- Kawamura, S. (1958). Matriarchal social ranks in the Minoo-B troop: a study of the rank system of Japanese monkeys (in Japanese). *Primates*, 1, 148-156. <https://doi.org/10.1007/BF01813701>

- Lawick-Goodall, J. van (1968). A preliminary report on expressive movements and communication in the Gombe Stream chimpanzees. En P. C. Jay (Ed.), *Primates: studies in adaptation and variability* (pp. 313-519). New York: Holt, Rinehart and Winston.
- Lefebvre, L. (1995). Culturally-transmitted feeding behaviour in primates: evidence for accelerating learning rates. *Primates*, 36, 227-239. <https://doi.org/10.1007/BF02381348>
- Matsuzawa, T. (2003). Koshima monkeys and Bossou chimpanzees: culture in nonhuman primates based on long-term research. En F. B. M. de Waal y P. L. Tyack (Eds.), *Animal social complexity: intelligence, culture, and individualized societies*. Cambridge, Mass.: Harvard University Press. <https://doi.org/10.4159/harvard.9780674419131.c28>
- Matsuzawa, T., Biro, D., Humle, T., Inoue-Nakamura, N., Tonooka, R. y Yamakoshi, G. (2001). Emergence of culture in wild chimpanzees: education by master-apprenticeship. En T. Matsuzawa (Ed.), *Primate origins of human cognition and behavior* (pp. 557-574). Tokyo/Berlin/Heidelberg: Springer. [https://doi.org/10.1007/978-4-431-09423-4\\_28](https://doi.org/10.1007/978-4-431-09423-4_28)
- Mitchell, R. W., Thompson, N. S. y Miles, H. L. (1997). *Anthropomorphism, anecdotes, and animals*. Albany, NY: SUNY Press.
- Nishida, T. (1968). The social group of wild chimpanzees in the Mahali Mountains. *Primates*, 9, 167-224. <https://doi.org/10.1007/BF01730971>
- Nishida, T. (1987). Local traditions and cultural transmission. En B. B. Smut, D. L. Chene, R. M. Seyfarth, R. W. Wrangham y T. T. Struhsaker (Eds.), *Primate societies* (pp. 462-474). Chicago: University of Chicago Press.
- Premack, D. y Premack, A. J. (1994). Why animals have neither culture nor history. En T. Ingold (Ed.), *Companion Encyclopedia of Anthropology* (pp. 350-365). London: Routledge.
- Rendell, L. y Whitehead, H. (2001). Culture in whales and dolphins. *Behav. Brain Sci.*, 24, 309-382. <https://doi.org/10.1017/S0140525X0100396X>
- Reynolds, V. (1992). The discovery of primate kinship systems by Japanese anthropologists. En N. Itoigawa, Y. Sugiyama, G. Sackett y R. Thompson (Eds.), *Topics in primatology* (vol 2, pp. 79-87). Tokyo: University of Tokyo Press.
- Sakura, O. (1998). Similarities and varieties: a brief sketch on the reception of Darwinism and sociobiology in Japan. *Biol. Philos.*, 13, 341-357. <https://doi.org/10.1023/A:1006504623820>
- Takasaki, H. (2000). Traditions of the Kyoto School of field primatology in Japan. En S. Strum y L. M. Fedigan (Eds.), *Primate encounters* (pp 151-164). Chicago: University of Chicago Press.

- Todes, D. (1989). *Darwin without Malthus: the struggle for existence in Russian evolutionary thought*. New York: Oxford University Press.
- Tomasello, M. (1994). The question of chimpanzee culture. En R. Wrangham et al. (Eds.), *Chimpanzee cultures* (pp. 301-317). Cambridge, Mass.: Harvard University Press.
- Waal F. B. M. de (1996). *Good natured*. Cambridge, Mass.: Harvard University Press.
- Waal, F. B. M. de (1998). No imitation without identification. *Behav. Brain Sci.*, 21, 689. <https://doi.org/10.1017/S0140525X98271743>
- Waal, F. B. M. de (1999). Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. *Philos. Topics*, 27, 255-280. <https://doi.org/10.5840/philtopics199927122>
- Waal, F. B. M. de (2001). *The ape and the sushi master*. New York: Basic Books.
- Waal, F. B. M. de y Tyack, P. L. (2003). Animal social complexity: intelligence, culture, and individualized societies. Cambridge, Mass: Harvard University Press.
- Watanabe, K. (1994). Precultural behavior of Japanese macaques: longitudinal studies of the Koshima troops. En R. A. Gardner, A. B. Chiarelli, B. T. Gardner y F. X. Plooj (Eds.), *The ethological roots of culture* (pp. 81-94). Dordrecht: Kluwer. [https://doi.org/10.1007/978-94-011-0998-7\\_5](https://doi.org/10.1007/978-94-011-0998-7_5)
- West, M. J., King, A. P. y White, D. J. (2003). Discovering culture in birds: the role of learning and development. En F. B. M. de Waal y P. L. Tyack (Eds.), *Animal social complexity: intelligence, culture, and individualized societies* (pp. 470-492). Cambridge, Mass.: Harvard University Press. <https://doi.org/10.4159/harvard.9780674419131.c36>
- Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W. y Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399, 682-685. <https://doi.org/10.1038/21415>
- Yamada, M. (1963). A study of blood-relationship in the natural society of the Japanese macaque – an analysis of co-feeding, grooming, and playmate relationships in Minoo-B-troop. *Primates*, 4, 43-65. <https://doi.org/10.1007/BF01733671>
- Yoshimi, K. (1998) Imanishi Kinji's biosociology as a forerunner of the semiosphere concept. *Semiotica*, 120, 273-297.



eISSN: 1989-3612

DOI: <https://doi.org/10.14201/art2024.31935>

# NORMATIVIDAD NATURAL: EL “ES” Y EL “DEBE” DEL COMPORTAMIENTO ANIMAL<sup>1</sup>

## TRADUCCIÓN

Frans B.M. DE WAAL

*Living Links, Yerkes National Primate Research Center,*

*Emory University, Atlanta, GA, United States*

dewaal@emory.edu

**RESUMEN:** Suele suceder que la evolución del comportamiento se considera irrelevante para la comprensión de la moral humana, por el hecho de que carece de carácter normativo (el ‘debe’) y consiste enteramente en descripciones de cómo son las cosas o cómo sucedieron (el ‘es’). No obstante, el comportamiento que es producto de la evolución, incluido el de otros animales, no está completamente desprovisto de normatividad. Si se define la normatividad como la adhesión a un ideal o estándar, existe amplia evidencia de que los animales tratan a sus relaciones sociales de esta manera. En otras palabras, persiguen valores sociales. En este artículo reviso la evidencia sobre el hecho de que los primates no humanos intentan activamente preservar la armonía dentro de su red social, por ejemplo, reconciliándose después de un conflicto, protestando contra las divisiones desiguales y deteniendo peleas. Al hacerlo, corrigen las desviaciones respecto de un estado ideal. Sumado a ello, y con el fin de

1. Traducción: E. Joaquín Suárez-Ruiz y Laura Danón. Esta traducción fue autorizada tanto por su autor, Frans de Waal, como por su editorial original, *Behaviour* (Brill).

prevenir tales desviaciones, muestran autocontrol emocional y resolución anticipada de conflictos. El reconocimiento de la orientación hacia una meta y el carácter normativo del comportamiento social animal nos permite cerrar parcialmente la brecha entre el 'es' y el 'debe' erigida en relación con el comportamiento moral humano.

*Palabras clave:* moral, normatividad, resolución de conflictos, interés comunitario, justicia, aversión a la inequidad, control emocional.

## 1. INTRODUCCIÓN

Uno de los problemas más irritantes a los que se enfrentan los intentos de fundamentar la moral en la biología es la llamada división 'ser/deber ser'. Si acaso se trata de una división real o no, depende en parte de cómo enunciemos la pregunta. Vale resaltar que quien la formulara inicialmente, David Hume (1739), no la vio, de hecho, como una división nítida. Hace ya casi tres siglos, lo que nos pidió fue que tuviéramos la precaución de no asumir que podemos derivar el 'debe' del 'es' y agregó que, si vamos a intentarlo, debemos dar una razón para ello. Habiendo notado la frecuencia con la que algunos autores pasaban de descripciones de cómo son las cosas a declaraciones sobre cómo deberían ser, agregó:

Este cambio es imperceptible, pero resulta, sin embargo, de la mayor importancia. En efecto, en cuanto que este 'debe' o "no debe" expresa alguna nueva relación o afirmación, es necesario que esta sea observada y explicada y que al mismo tiempo se dé razón de algo que parece completamente inconcebible, a saber: cómo es posible que esta nueva relación se deduzca de otras totalmente diferentes (Hume, 1739, 335).<sup>2</sup>

En otras palabras, la forma en que nos parece que deben comportarse los humanos no es simplemente un reflejo de la naturaleza humana. Así como no se pueden inferir las normas de tránsito de la descripción de un automóvil, tampoco se pueden inferir códigos morales a partir de saber quiénes o qué somos. El punto de Hume se entiende bien, pero está muy lejos de la ampliación que hicieron de él algunos filósofos posteriores, quienes convirtieron su pedido de cautela en la "guillotina de Hume", proclamando un abismo infranqueable entre el 'es' y el 'debe' (Black, 1970). De ninguna manera existe acuerdo sobre este tema (sigue siendo un debate

2. N. de los T.: Se utilizó la traducción de Félix Duque. Hume, D. 1992 (1739). *Tratado De La Naturaleza Humana*. Tecnos.

filosófico perenne), pero algunos han ido tan lejos empuñando esta guillotina que pretendieron acabar con todos y cada uno de los intentos, incluso los más cautelosos, de aplicar la lógica evolutiva o la neurociencia a la comprensión de la moral humana. La ciencia no puede decirnos cómo entender la moral, argumentan. Esto bien puede ser cierto, pero la ciencia sí puede ayudar a explicar por qué ciertos resultados son preferidos/favorecidos por sobre otros y, por lo tanto, por qué la moralidad es como es. Para empezar, no tendría sentido diseñar reglas morales que fueran imposibles de seguir, al igual que no tendría sentido hacer reglas de tránsito que los automovilistas no puedan obedecer. Esto se conoce como el argumento "el 'debe' implica el 'puede'". La moral debe amoldarse a la especie a la que está destinada.

El 'es' y el 'debe' son como el yin y el yang de la moralidad. Tenemos ambos, necesitamos ambos, no son lo mismo, pero no se pueden desvincular por completo. Se complementan entre sí (ver también Kitcher, 2014, en este número<sup>3</sup>). Al enfatizar la importancia de la naturaleza humana, el mismo Hume (1739) ignoró la 'guillotina' que lleva su nombre: vio la moral como un producto de las emociones, colocando la empatía (a la que llamó 'simpatía') en la parte superior de su lista. Esta opinión no representaba ninguna contradicción, ya que, como se mencionó más arriba, lo único que sugería era que tengamos prudencia en el paso del cómo somos al cómo debemos comportarnos (Baier, 1991). Nunca dijo que tal movimiento estuviera prohibido (aunque podría no haber estado de acuerdo con Singer [1973], para quien el debate sobre la división entre el 'es' y el 'debe' es una "trivialidad" que depende por completo de la definición de moral).

Si bien estoy de acuerdo con muchos filósofos en que es difícil, tal vez incluso imposible, razonar desde el nivel de cómo son las cosas hasta el cómo deberían ser, exploraré aquí si acaso la brecha divisoria resulta igualmente extensa cuando dejamos el dominio conceptual y entramos en el de las tendencias y motivaciones conductuales reales. ¿Qué tal si la moral no se construye racionalmente, sino que, como pensaba Hume, se basa en valores emocionales? ¿Qué tal si la biología no está sólo en el lado del 'es' de la ecuación, sino que también nos informa sobre el lado del 'debe', por ejemplo, al explicar qué valores perseguimos y por qué razón evolutiva? Cada organismo se esfuerza por lograr ciertos resultados. La supervivencia es uno, la reproducción es otro, pero muchos organismos

3. N. de los T.: de Waal se refiere al siguiente texto: Kitcher, P. (2014). Is a naturalized ethics possible? *Behaviour*, 151(2-3), 245-260.

también persiguen efectos sociales que se acercan a los respaldados por la moral humana.

El hecho de que el comportamiento animal no está libre de normatividad (definida como la adhesión a un ideal o estándar) apenas necesita argumentación. Tengamos en cuenta, por ejemplo, la reacción de la araña ante una red dañada. Si el daño es extenso, abandonará su telaraña, pero la mayoría de las veces entrará en modo de reparación: devolverá la telaraña a su estado funcional anterior, rellenando agujeros o ajustando los hilos dañados mediante otros nuevos (Eberhard, 1972). Del mismo modo, perturbar un nido de hormigas o un montículo de termitas conduce a una reparación inmediata, al igual que los daños generados a un dique de castores o a un nido de pájaros. La naturaleza está llena de estructuras físicas construidas por animales guiados por un patrón que indica cómo debería verse la estructura. Dicho patrón motiva la reparación o el ajuste tan pronto como la estructura se desvía del ideal. En otras palabras, los animales tratan estas estructuras de manera normativa. No estoy pensando necesariamente aquí en un juicio normativo. No está claro si los animales sienten la obligación de comportarse de una manera particular, ni asumo que cada individuo de una gran colonia tiene noción del nido como un todo, pero es innegable que los animales persiguen objetivos, colectiva o individualmente.

La cuestión aquí es si hacen lo mismo con respecto a las relaciones sociales y a la sociedad en general ¿Buscan ciertos resultados sociales y corrigen o desalientan las desviaciones de las expectativas? ¿Adoptan un enfoque normativo de las relaciones sociales y, de ser así, está guiado por el mismo tipo de emociones y valores que subyacen a la moral humana? Churchland (2011, 175) sugiere que hay una transición de las emociones sociales a los valores morales, sosteniendo que "las emociones básicas son la forma en que la Madre Naturaleza nos orienta para hacer lo que prudentemente debemos". La pregunta es si este mismo movimiento es reconocible en otras especies.

## 2. JERARQUÍA SOCIAL Y CONTROL DE IMPULSOS

Lo opuesto a la moral es que simplemente hagamos "lo que queremos", siendo un supuesto subyacente que lo que queremos no es lo moralmente bueno. Este sigue siendo un argumento religioso frecuente en contra de la ética naturalizada (Gallagher, 2004). Desde este punto de vista, la moral se basa en la capacidad exclusivamente humana de inhibir las tendencias naturales (Huxley, 1894). Por ejemplo, Kitcher (2006)

etiquetó a los chimpancés como "caprichosos" [*wantons*], definiéndolos como criaturas vulnerables a cualquier impulso que incida sobre ellos. En algún momento de nuestra evolución superamos esta impulsividad caprichosa y es esa superación, justamente, la que nos hizo humanos. Según Kitcher (2006, 136), este proceso comenzó con la "consciencia de que ciertos tipos de comportamientos proyectados podrían tener resultados problemáticos".

Sin embargo, innumerables animales viven con un tipo similar de conocimiento, no sólo cuando intentan evitar ser detectados por depredadores o presas mediante la supresión del sonido y el movimiento, sino también en el ámbito social. Una jerarquía de dominancia es un sistema gigante de inhibiciones sociales, que, sin duda, allanó el camino para la moral humana, el cual también es un sistema de este tipo. El control de los impulsos es clave para evitar "resultados problemáticos". En macacos y otros primates, los machos de bajo rango varían su comportamiento dependiendo de la presencia o ausencia del macho alfa. Tan pronto como el alfa les da la espalda, los otros machos se acercan a las hembras. En un experimento que ponía a prueba este principio, los machos de bajo rango se negaban a acercarse a las hembras mientras el dominante miraba desde el interior de una caja transparente, pero tan pronto como este macho era retirado, copulaban libremente con ellas. Los machos de bajo rango también aprovechaban la ocasión para realizar las típicas demostraciones de vigor [*bouncing displays*] de los machos de alto estatus. Después de tales episodios, sin embargo, estaban excesivamente nerviosos al reunirse con el macho alfa, mostrando los dientes con tanta sumisión al saludarlo que los experimentadores interpretaron su comportamiento como un reconocimiento implícito de que habían violado un código social (Coe y Roseblum, 1984). Tal vez las reglas sociales no se obedezcan simplemente en presencia de los dominantes y se olviden en su ausencia, sino que hasta cierto punto también se interioricen. Ahora bien, cuando los científicos han tratado de medir el grado de internalización de las reglas impuestas por humanos en los perros, estudiando su comportamiento de aspecto culpable después de las violaciones a las mismas, no han encontrado mucho más que cierto efecto directo del comportamiento del dueño sobre el perro (Vollmer, 1977; Horowitz, 2009).

No sólo los individuos de bajo rango, sino también los de alto rango, se benefician del control de impulsos. Por ejemplo, un chimpancé macho alfa experimentado (*Pan troglodytes*) puede ser desafiado de modo evidente por parte de un macho más joven, que arroja piedras en su dirección o hace una llamativa exhibición de fuerza con todos los pelos de punta (este es un modo de poner a prueba los nervios del alfa) y, sin embargo, ignorar

totalmente el estruendo, como si apenas se diera cuenta, lo que obliga a su retador o a rendirse o a intensificar aún más la situación (de Waal, 1982).

En última instancia, las inhibiciones asociadas con la jerarquía llegan a través del castigo. Por ejemplo, después de haber privado de agua a una gran manada de monos rhesus (*Macaca mulatta*) durante tres horas, se puso a su disposición un único estanque lleno de agua. Todos los adultos fueron a beber en orden jerárquico. No obstante, los infantes y los jóvenes bebían con los machos de más alto rango y se mezclaban con la línea matricial superior, ignorando así la jerarquía social. Es recién en el tercer año de vida, a través de exclusiones y castigos cada vez mayores, que los menores comienzan a aprender su lugar en el orden de rango general y a converger así con el rango de su madre (de Waal, 1993).

En el caso de los simios, dado que se desarrollan más lentamente que los monos, los jóvenes andan prácticamente impunes durante sus primeros cuatro años de vida. Por ejemplo, usan la espalda de un macho dominante como trampolín, roban comida de las manos de otros o golpean a un joven mayor tan fuerte como pueden: nada de lo que hacen resulta/parece incorrecto. Uno puede imaginar la conmoción de un joven cuando es rechazado o castigado por primera vez. Los castigos más dramáticos son los de aquellos varones jóvenes que se aventuran demasiado cerca de una hembra sexualmente atractiva (de Waal, 1982; Figura 1). Los machos jóvenes sólo necesitan una o dos lecciones de este tipo. A partir de allí, todo macho adulto puede hacer que salten lejos de una hembra simplemente con una mirada o un paso adelante. De esta manera, los más jóvenes aprenden a controlar sus impulsos sexuales o, por lo menos, se vuelven más discretos a la hora de actuar cuando son guiados por ellos.

La capacidad para controlar los impulsos puede testearse experimentalmente de la misma manera que se testea la demora de la gratificación en niños (Mischel *et al.*, 1972; Logue, 1988). A los niños se les da un malvavisco con la promesa de que, si no se lo comen, tendrán otro más. Muchos niños tienen la capacidad de esperar durante minutos. De manera similar, tanto los simios (Beran *et al.*, 1999) como los monos (Amici *et al.*, 2008) dejan pasar una recompensa inmediata en favor de una mejor y más tardía. También se ha mostrado que los chimpancés, como los niños, juegan más con sus juguetes en presencia de recompensas que se acumulan, lo que sugiere que intentan auto-distraerse frente a la tentación. Esta acción permite a los simios demorar la gratificación hasta 18 minutos (Evans y Beran, 2007). Otros estudios han demostrado que los simios pueden anular un impulso inmediato en favor de necesidades futuras, un aspecto esencial para una planificación exitosa de la acción (Osvath y Osvath, 2008).



**Figura 1.** Un macho joven, de unos 4 años, ha mostrado demasiado interés por una de las hembras en celo y ahora está siendo castigado por un macho adulto, que le ha metido el pie en la boca y lo balancea. Esto servirá como una lección para el resto de la vida del joven macho sobre la competitividad de los machos en torno a las hembras sexualmente atractivas. Fotografía de Frans de Waal

El mismo entrelazamiento entre emoción y cognición reconocido en humanos parece estar presente en nuestros parientes cercanos, incluyendo el control deliberado de las emociones. En la medida en que dicho control está mediado por los lóbulos frontales, es importante señalar que la opinión popular de que esta parte del cerebro está excepcionalmente desarrollada en nuestra especie es errónea. El cerebro humano es esencialmente un cerebro de mono ampliado linealmente (Herculano-Houzel, 2009; Barton y Venditti, 2013).

### 3. NORMATIVIDAD UNO A UNO

La moral puede ser definida como un sistema de reglas que gira en torno a ayudar o, al menos, a no dañar a los demás seres humanos. Su propósito es el bienestar de los demás y, a menudo, antepone los intereses de la comunidad frente a los del individuo. No niega el interés propio, pero pone restricciones a su búsqueda con el fin de promover una sociedad cooperativa (de Waal, 1996, 2006). Esta definición funcional diferencia a la moral de las costumbres y los hábitos, como comer con cuchillo y tenedor versus comer con palillos o con las manos desnudas. La distinción entre reglas morales y convenciones ya se ve claramente en niños pequeños (Killen y Rizzo, 2014, este número). En otro trabajo he distinguido dos niveles de reglas morales: (a) reglas en el nivel uno a uno (diádicas) de las relaciones sociales, y (b) reglas en el nivel comunitario (de Waal, 2013). La Tabla 1 resume ejemplos en el nivel uno a uno.

Ideal	Desviación	Reparación o corrección	Restauración
Jerarquía	Desobediencia o desafío al rango	Castigo o restablecimiento de dominancia	Annonía
Relación cercana	Conflicto	Reconciliación	Armonía
Cooperación	Recompensa desigual	Protesta o compartición	Annonía
Juego relajado	Compañero lastimado	Señales de remediación	Annonía

**Tabla 1.** Cuando los individuos buscan preservar relaciones sociales armoniosas, aplican la normatividad uno a uno. Su comportamiento refleja el valor que atribuyen a las buenas relaciones. Esta tabla ofrece cuatro ejemplos: restauración de la jerarquía de dominancia, reparación de relaciones, reacciones negativas a la inequidad y reanudación del juego. En todos los casos, los primates y otros animales buscan activamente que una relación social retorne a su estado original

#### 3.1. Reconciliación

El nivel uno a uno gira en torno a la preservación de relaciones valiosas. Una de sus expresiones más comunes es la resolución de conflictos, reportada por primera vez por de Waal y van Roosmalen (1979). Un ejemplo típico refiere a dos chimpancés machos que luego de haber estado persiguiéndose y gritándose, descansan en un árbol (Figura 2). Diez minutos después, un macho extiende su mano al otro, rogándole un abrazo. En cuestión de segundos, se abrazan, se besan y bajan juntos al suelo para acicalarse. Entendiéndolo como una reconciliación, este proceso se define como un contacto amistoso que ocurre poco después de un



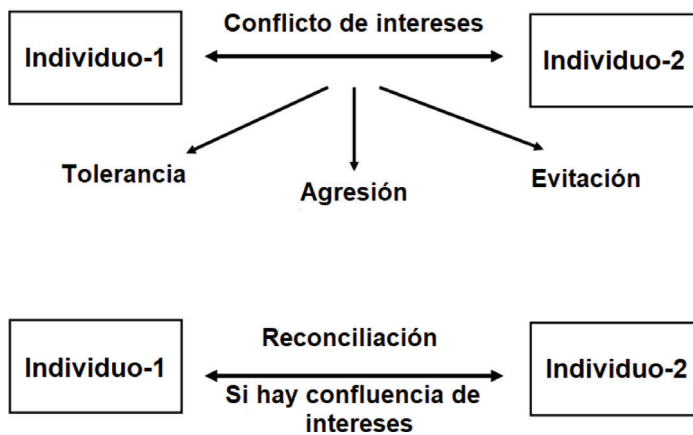


**Figura 2.** La situación después de un prolongado y ruidoso conflicto entre dos machos adultos en un zoológico. El macho desafiado (izquierda) había huido hacia el árbol. 10 minutos después su oponente le extendió una mano. En cuestión de segundos, los dos machos mantuvieron un reencuentro físico. Fotografía de Frans de Waal

conflicto entre dos partes. Mientras que el besarse es la forma más típica de reconciliación entre chimpancés, los bonobos lo hacen mediante el comportamiento sexual (de Waal, 1987) y los macacos rabones llevan a cabo un comportamiento ritual en el cual esperan hasta que el subordinado se presente para luego sostener sus caderas (de Waal y Ren, 1988). Cada especie tiene su propio camino, pero el principio básico sigue siendo el mismo: los oponentes vuelven a reunirse después de una pelea.

La primatología se ha interesado durante mucho tiempo en las relaciones sociales, por lo que la idea de reparación de relaciones, implícita en la etiqueta de reconciliación, atrajo rápidamente su atención. Ahora sabemos que una treintena de especies diferentes de primates se reconcilian después de las peleas y que esa reconciliación no se limita a los primates. Hay pruebas de este mecanismo en hienas, delfines, lobos, cabras domésticas, etc. La razón por la que la reconciliación está tan extendida es que restituye relaciones que han sido dañadas por la agresión pero que, sin embargo, son esenciales para la supervivencia. Dado que muchos animales establecen relaciones de cooperación en las que ocasionalmente surgen conflictos, los mecanismos de reparación son esenciales. El creciente campo de la resolución de conflictos en animales ha sido examinado por de Waal (2000) y Aureli y de Waal (2000).

La mayoría de estos estudios apoyan la Hipótesis de la Relación Valiosa, que puede formularse de la siguiente manera: "La reconciliación ocurrirá especialmente entre individuos que tienen mucho que perder si su relación se deteriora". Esta hipótesis también ha sido respaldada por un elegante experimento que manipuló el valor de la relación, promoviendo la cooperación entre monos y aumentando de este modo su disposición a reconciliarse después de las peleas (Cords y Thurnheer, 1993). Las ideas anteriores se han formalizado en el Modelo Relacional, que sitúa el conflicto en un contexto social. La agresión es vista como una de las varias opciones para resolver un conflicto de intereses. Otras opciones son evitar al adversario (común en especies jerárquicas y territoriales) y compartir recursos (común en especies tolerantes). Sopesando los costos y beneficios de cada opción, el conflicto puede escalar hasta el punto de la agresión, después de lo cual todavía existe la opción de reparar el daño por medio de la reconciliación, la cual suele verse favorecida cuando las partes tienen intereses superpuestos (de Waal, 2000; Figura 3). Aplicando la misma metodología primatológica estandarizada a niños humanos, se han obtenido resultados muy similares (Verbeek *et al.*, 2000).



**Figura 3.** De acuerdo con el modelo relacional, el comportamiento agresivo es una de las varias formas en que se pueden resolver los conflictos de intereses. Otras formas posibles son la tolerancia (p. ej., compartir recursos) y evitar la confrontación (p. ej., por parte de los subordinados a los dominantes). Si ocurre una agresión, dependerá de la naturaleza de la relación social el que se lleven a cabo o no intentos de reparación. En el caso de que exista un fuerte interés mutuo en mantener la relación, es probable que haya una reconciliación. Las partes negociarán los términos de su relación pasando por ciclos de conflicto y reconciliación. Véase de Waal (2000)

### 3.2. Resolución preventiva de conflictos

El hecho de que los primates se protegen contra los efectos perjudiciales del conflicto y la angustia es algo que resulta visible durante el juego. Cuando los jóvenes tienen edades muy diferentes, los juegos a menudo se vuelven demasiado brutos para el compañero más joven, como cuando tuercen sus piernas o cuando un mordisqueo se convierte en un tarascón. A la menor señal de peligro, su madre interrumpirá el juego. Normalmente, excepto por las roncadas y jadeantes risas de los simios (semejantes a la risa humana), el juego es completamente silencioso (van Hooff, 1972). Tras registrar cientos de estas escenas de lucha descubrimos que los chimpancés jóvenes emiten este tipo de vocalización especialmente cuando la madre de un compañero de juegos más joven está mirando. Es decir, lo hacen más en presencia de la madre que cuando están solamente con el chimpancé más joven. Puede que las vocalizaciones en cuestión tengan la función de evitar la intervención materna, tranquilizándola sobre la naturaleza benigna de la interacción (Flack *et al.*, 2004).

Por su parte, Bekoff (2001) analizó videos de juegos entre perros, lobos y coyotes, concluyendo que el juego de los cánidos está sujeto a reglas, genera confianza, requiere considerar al otro y enseña a los jóvenes a comportarse. El altamente estereotipado comportamiento de "reverencia de juego" (un animal se agacha sobre sus extremidades anteriores mientras levanta su trasero), ayuda a diferenciar el juego del sexo o el conflicto, con los cuales puede llegar a confundirse. El juego cesa abruptamente, sin embargo, tan pronto como un compañero hiere a otro. El transgresor precisará realizar una nueva reverencia, después de lo cual el compañero podría continuar el juego. Bekoff establece un paralelo con la moral:

Durante el juego social, mientras los individuos se divierten en un entorno relativamente seguro, aprenden reglas básicas que resultan aceptables para los demás: qué tan fuerte pueden morder, qué tan bruscamente pueden interactuar y cómo resolver conflictos. Es muy importante jugar de manera justa y confiar en que los demás también lo harán. Hay códigos de conducta social que regulan lo que está permitido y lo que no, y la existencia de estos códigos podría tener algo que decir sobre la evolución de la moral (Bekoff, 2001, 85).

El comportamiento destinado a la preservación de las buenas relaciones es un indicio del gran valor que se le otorga a la armonía social. Kummer (1995) ofrece observaciones sorprendentes de la forma en que los líderes de harén de los babuinos hamadryas (*Papio hamadryas*),

al encontrarse frente a un árbol frutal demasiado pequeño como para alimentar a sus dos familias, suspenderán su inevitable confrontación huyendo literalmente el uno del otro, seguidos por sus respectivas hembras y crías. Los machos de chimpancé enfrentan un dilema similar. Varios de ellos pueden sentarse cerca de una hembra que anuncia su fertilidad con los genitales hinchados. En lugar de competir, los machos mantienen activamente la paz. Aunque con frecuencia echan un vistazo a la hembra, pasan el día acicalándose mutuamente. Recién cuando todos están lo suficientemente relajados, uno de ellos intentará aparearse (de Waal, 1982).

Si bien las descripciones anteriores son cualitativas, también se han cuantificado las técnicas de prevención de conflictos. Tras una sugerencia inicial de de Waal (1987), según la cual habría un pico de acicalamiento entre los bonobos en cautiverio justo antes de la hora de comer, precediendo la posible competencia y tensión, los estudios han procurado medir el comportamiento en torno a la llegada de alimentos, que en la mayoría de los zoológicos e instituciones ocurre en un momento predecible del día. Los chimpancés se acicalan más mientras esperan comida y participan en "celebraciones" marcadas por altos niveles de contacto corporal tranquilizador una vez que llega la comida (de Waal, 1992a; Koyama y Dunbar, 1996). Los bonobos, por otro lado, muestran un aumento en su comportamiento de juego antes de la comida y una gran cantidad de contacto sociosexual a su llegada (de Waal, 1987; Palagi *et al.*, 2006). Por lo tanto, los primates pueden anticipar la competencia por la comida y trabajan activamente para reducirla.

### 3.3. *Persiguiendo una división justa de recompensas*

Las reacciones negativas a las distribuciones de recompensas sesgadas, también conocidas como aversión a la inequidad (AI), son otra parte del mantenimiento de las relaciones diádicas. Los animales cooperativos precisan vigilar qué beneficios obtienen en relación con sus socios cooperativos, a fin de evitar que se aprovechen de ellos. En ausencia de una distribución equitativa, la cooperación mutualista podría convertirse fácilmente en una forma de altruismo por parte de quienes ganan menos. Este 'problema de resultado' [*outcome problem*] ha sido reconocido en humanos (Fehr y Schmidt, 1999), y es cada vez más un tema recurrente en la investigación con animales (Brosnan, 2011).

Los monos capuchinos son tan sensibles a la inequidad que ante las recompensas que se encuentran amontonadas espacialmente

(resultando por ello monopolizables por las partes dominantes) se reducen más sus tendencias cooperativas que en el caso de recompensas dispersas (de Waal y Davis, 2003). Brosnan y de Waal (2003) probaron la AI en un experimento simple en el que dos monos recibían recompensas iguales o desiguales por la misma tarea, por ejemplo, uno recibía rodajas de pepino y el otro uvas (preferidas por ellos). Los autores encontraron que, cuando ambos recibían lo mismo, los individuos que obtenían la recompensa menor no se veían afectados, pero si su compañero recibía un mejor trato, a menudo se negaban a realizar o aceptar la recompensa. Se encontraron resultados similares en chimpancés (Brosnan *et al.*, 2005). Sin embargo, las replicaciones experimentales que no requerían que los participantes desempeñasen una tarea, no dieron los mismos resultados (Braüer *et al.*, 2006; Roma *et al.*, 2006) (incluso en el caso de un estudio con los mismos monos que en el estudio original) (Dindo y de Waal, 2006). Por lo tanto, como podría predecirse desde una explicación evolutiva centrada en el desempeño de tareas y en la cooperación, las recompensas desiguales provocan reacciones negativas sólo en el contexto de una tarea que requiere esfuerzo. Finalmente, van Wolkenten *et al.* (2007) demostraron que las respuestas a la inequidad son realmente sociales en el sentido de que no pueden explicarse como reacciones negativas a recompensas menores cuando otras mejores están visibles. Es decir, la mera visibilidad tuvo poco efecto: las reacciones negativas ocurrieron sólo si las mejores recompensas eran consumidas por un compañero.

Se han observado respuestas AI similares en otras especies, tanto primates como no primates (Brosnan, 2011; Price y Brosnan, 2012; Range *et al.*, 2012). Sin embargo, existe una restricción: hasta ahora, la mayoría de los estudios sólo se refieren a la AI por parte del individuo que recibe menos, conocida como AI desventajosa. En la AI ventajosa, en cambio, los sujetos responden negativamente al recibir una recompensa mayor que la de su compañero. Los humanos muestran tanto la última respuesta como la primera. Brosnan y de Waal (2012) especulan que la AI ventajosa, que indica un pleno sentido de justicia, ocurre cuando los individuos anticipan las implicaciones negativas de la AI desventajosa en los demás. Con el fin de proteger la relación frente a los efectos erosivos de las tensiones surgidas cuando un individuo recibe menos que el otro, el que recibe más intenta evitarlo igualando la recompensa. Los autores señalan esto como un sentido de justicia de segundo orden: "Para prevenir conflictos al interior de relaciones cercanas o beneficiosas, el individuo aventajado se beneficiará de protestar o rectificar la situación" (Brosnan y de Waal, 2012, 341).

Hasta el momento no hay signos de justicia de segundo orden en monos, como por ejemplo en los monos capuchinos del estudio original. En los simios, sin embargo, la evidencia va en aumento. La primera señal provino de un estudio de Brosnan *et al.* (2010) en chimpancés, en el que no sólo se negaban regularmente a realizar o aceptar sus recompensas los participantes que recibían la menor recompensa, sino también respondían de este modo quienes que recibían la mejor recompensa. En otras palabras, cualquier inequidad, no sólo la desventajosa, era resistida. Tenía sentido, por tanto, testear a los chimpancés con el Juego del Ultimátum (JU), que es el estándar de oro del sentido de justicia humano. En el JU, un individuo (el Proponente) puede dividir el dinero con otro individuo (el Respondedor). Si el Respondedor acepta la oferta, ambos jugadores son recompensados con la división acordada. Sin embargo, si el Respondedor rechaza la oferta, ninguno de los jugadores es recompensado. En las culturas occidentales, las personas suelen ofrecer alrededor del 50% de la cantidad disponible (Guth, 1995; Camerer y Lowenstein, 2004), al igual que en la mayoría de las otras culturas (Henrich *et al.*, 2001). En contraste, un estudio del JU en chimpancés encontró que comparten la menor cantidad posible con los demás (Jensen *et al.*, 2006). Sin embargo, la metodología de este experimento se desvió sustancialmente del JU humano típico y no estaba claro si los simios entendieron completamente la tarea.

Para superar estas objeciones, Proctor *et al.* (2013) diseñaron un procedimiento del JU que resultaba más intuitivo tanto para chimpancés como para niños humanos de 3 a 5 años. A los Proponentes se les presentó la posibilidad de elegir entre dos fichas de diferente color que podrían intercambiar por comida con un experimentador humano. Mientras que un color representaba una distribución equitativa de la recompensa (3 rodajas de plátano versus otras 3), el otro representaba una distribución desigual a favor del Proponente (5 rodajas de plátano versus 1). El Proponente tenía que entregar la ficha a su socio, el Respondedor, sentado detrás de una reja de malla. Quienes respondían podían aceptar la ficha y devolvérsela al experimentador, o rechazarla no devolviendo la ficha. Como en el JU humano típico, los Proponentes necesitaban la colaboración del Respondedor.

Las elecciones de fichas se compararon con elecciones en presencia de Respondedores que carecían de cualquier influencia. Los chimpancés resultaron sensibles a las contingencias del juego, de la misma forma que los humanos. Si su compañero tenía el control, dividían las recompensas en partes iguales con mayor frecuencia. Sin embargo, en ausencia de la influencia de su compañero, preferían la opción que les otorgaba

la mayor proporción de recompensa. Siendo que los niños se comportan de manera similar, el estudio sugiere que los humanos y los chimpancés comparten patrones de toma de decisiones proactivas en relación con los resultados justos (Proctor *et al.*, 2013).

#### 4. PREOCUPACIÓN COMUNITARIA

En comparación con lo que ocurre en el caso de la normatividad individual uno a uno, hay muchos menos signos de normatividad a nivel comunitario en primates no humanos. Este es el nivel en que la moral humana parece ser única, ya que habitualmente extendemos nuestro razonamiento moral hacia la sociedad en su conjunto, especulando qué le sucedería a nuestra comunidad si acaso todos actuáramos de cierta manera. De hecho, solemos extender nuestro sistema de valores a interacciones en las que no estamos involucrados de manera directa. Siguiendo a Westermarck (1917, 238), las emociones morales difieren de las ordinarias "por su desinterés, aparente imparcialidad y generalidad". Mientras que las emociones típicas conciernen sólo a nuestros intereses personales — cómo nos han tratado o cómo queremos que nos traten—, las emociones morales van más allá. Se ocupan del bien y del mal en un nivel más abstracto. Sólo cuando realizamos juicios respecto de cómo debería ser tratado alguien bajo ciertas circunstancias, estamos hablando de un juicio moral. Para transmitir la misma idea, Smith (1759) nos pidió que imagináramos cómo un "espectador imparcial" juzgaría el comportamiento humano.

Esto no quiere decir que el nivel de la normatividad comunitaria esté completamente ausente del comportamiento de nuestros parientes cercanos. En un trabajo anterior he denominado este nivel como el de la "preocupación comunitaria" (de Waal, 1996). Existen muchos ejemplos de vigilancia y mediación imparciales que parecen reflejar valores a nivel de la comunidad. En algunas especies, las intervenciones de los miembros de rango más alto ponen fin a las peleas o al menos reducen la gravedad de la agresión. A menudo, los chimpancés machos de alto rango juegan este papel en peleas entre hembras y/o entre jóvenes de su grupo (de Waal, 1982). Por ejemplo, si dos jóvenes están jugando y se desata una pelea, el macho alfa puede acercarse al área del conflicto para detener la pelea. Al suspenderla antes de que las madres intervengan y comiencen a pelear entre ellas, reduce los niveles de agresión dentro del grupo y evita también que la pelea de los jóvenes se intensifique.

Este patrón de comportamiento se conoce como "rol de control" (cf. Bernstein y Sharpe, 1966). De Waal (1982) proporcionó descripciones y análisis detallados, junto con datos que muestran que, durante la adopción de dicho rol, los machos ignoran sus lazos sociales con los participantes del conflicto. Mientras que la mayoría de los individuos apoyan a sus amigos y parientes, los machos controladores intervienen independientemente de sus preferencias sociales habituales (de Waal, 1992b). La capacidad de dejar de lado tales preferencias sugiere una forma rudimentaria de justicia en los sistemas sociales de los primates no humanos. También se ha observado vigilancia imparcial en chimpancés salvajes (Boehm, 1994), y un estudio reciente que comparó este comportamiento en varios grupos cautivos concluyó que estabiliza la dinámica social (von Rohr *et al.*, 2012). Un estudio experimental de Flack *et al.* (2005), en el que se separó temporalmente del resto del grupo a miembros clave en lo que al control se refiere, mostró su importancia para el mantenimiento del aseo personal, el juego y otros signos de una sociedad armoniosa.

Otro método importante de resolución de conflictos que se ha identificado en grupos de primates es la mediación. La mediación ocurre cuando, en medio de un conflicto, un tercero se convierte en el puente entre dos oponentes incapaces de reconciliarse sin ayuda externa. La misma es caracterizada en el siguiente ejemplo (de Waal y van Roosmalen, 1979):

A veces, especialmente después de conflictos graves entre dos machos adultos, los dos oponentes eran reunidos por una hembra adulta. La hembra se acercaba a uno de los machos, lo besaba, lo tocaba o se presentaba ante él y luego, si el macho en cuestión la seguía, caminaba lentamente hacia el otro macho. El primero la seguía muy de cerca (a menudo inspeccionando sus genitales) y sin mirar al otro. En algunas ocasiones, la hembra miraba hacia atrás a su seguidor y, a veces, volvía hacia el que se había quedado atrás tirando de su brazo para hacer que la siguiera. Luego de que la hembra se sentaba cerca del otro macho, ambos comenzaban a acicalarla y simplemente continuaban acicalándose entre sí cuando ella se iba.

Llamar a tal comportamiento una expresión de interés comunitario de ninguna manera implica que no haya beneficios para el ejecutante. En los animales que viven en sociedad existe una gran superposición entre intereses comunitarios e individuales, de allí que cada individuo seguramente tenga interés en que su comunidad alcance cierto nivel de armonía y cooperación. Por lo tanto, el término 'preocupación comunitaria' no implica



ningún sacrificio y, menos aún, una selección a nivel grupal. Simplemente establece que los individuos pueden promover los intereses de su comunidad como un todo, lo que bien puede ser para su propio beneficio al mismo tiempo que beneficia a otros.

Finalmente, el prestigio y la reputación son una parte fundamental del por qué los seres humanos a menudo actúan en nombre de la comunidad, incluso cuando no se benefician directamente de ello. En los simios pueden verse algunos destellos de reputación. Por ejemplo, si una pelea importante se sale de control, los espectadores van a despertar al macho alfa dándole algunos golpecitos. Conocido como el árbitro más eficaz, se le insta a intervenir. También prestan atención a cómo un individuo trata a otro, como sucede en los experimentos en los que prefieren interactuar con un humano que ha mostrado una actitud positiva hacia los demás, por ejemplo, el haber compartido comida con otros simios (Russell *et al.*, 2008; Subiaul *et al.*, 2008; Herrmann *et al.*, 2013). En nuestros estudios, descubrimos lo siguiente: cuando dejábamos a la colonia observar a dos chimpancés que utilizaban trucos diferentes pero igualmente simples para obtener recompensas, preferían seguir al modelo de estatus superior. Este comportamiento evidenciaba el denominado 'efecto de prestigio', según el cual los simios imitan preferentemente a los miembros destacados de su comunidad (Horner *et al.*, 2010).

Esta evidencia sugiere que los chimpancés realizan acciones que benefician a la comunidad en su conjunto y, a su vez, que poseen reputaciones individuales relacionadas con cómo ciertos individuos tratan a los demás o cuán dignos de imitación son. No obstante, dichos comportamientos están todavía muy lejos del interés humano por los estándares de la comunidad y por el bienestar del conjunto. Es especialmente en el nivel del interés comunitario y de la construcción de la reputación que los sistemas morales humanos se desvían de la normatividad hallada en otros primates.

## 5. CONCLUSIÓN

A la luz de lo desarrollado y en relación con la división entre el 'es' y el 'debe', la postura que supone que la biología, incluido el comportamiento animal, reside completamente en el lado del 'es', resulta difícil de mantener. Obviamente, podemos describir el comportamiento animal dejando de lado todas y cada una de las referencias a objetivos, intenciones y

valores (del mismo modo que podemos describir el comportamiento humano de esta manera), pero tales descripciones pasan por alto un aspecto básico: los primates no humanos, así como también muchos otros animales, persiguen resultados específicos, tanto en relación con la construcción de ciertas estructuras físicas, por ejemplo nidos y redes, como en relación con los vínculos sociales. Intentan activamente preservar la armonía dentro de su comunidad social y con frecuencia corrigen las desviaciones de este ideal, por ejemplo, reconciliándose después de un conflicto, protestando contra divisiones desiguales o deteniendo peleas. Al corregir o tratar de corregir las desviaciones de un estado ideal, se comportan normativamente. A su vez, con el fin de prevenir dichas desviaciones, muestran autocontrol emocional y resolución anticipada de conflictos. Esto conlleva que el paso del comportamiento de los primates a las normas morales humanas sea un salto menor de lo que comúnmente se supuso.

Sin embargo, es probable que siga habiendo diferencias. Los otros primates parecen no extender las normas más allá de su entorno social inmediato y parecen no preocuparse por las relaciones sociales o situaciones en las que no participan de manera directa. A diferencia de los humanos, es posible que no sientan ninguna obligación de ser buenos ni que experimenten culpa y vergüenza cuando se equivocan. No sabemos si otros animales experimentan tales sentimientos vinculados con el 'debe'. Es posible argumentar que, si bien su comportamiento podría entenderse como normativo en el sentido de que busca ciertos resultados, lograrían llevarlos a cabo sin la necesidad de realizar un juicio normativo. Es decir, podrían evaluar el comportamiento social como exitoso o no en relación con la consecución de sus objetivos, pero no en términos de lo correcto y lo incorrecto. Por otro lado, su comportamiento a veces sugiere una suerte de evaluación de acciones pasadas, como cuando un bonobo muerde a otro y poco después se le acerca, recordando el lugar exacto de la mordedura, sólo para pasar media hora lamiéndole la herida infligida ( de Waal, 1989). Sin embargo, dada la inaccesibilidad de la experiencia animal, la existencia de una normatividad internalizada sigue siendo altamente especulativa. Por el momento, este artículo hace una afirmación más débil: en la medida en que el 'debe' de la moral humana refleja una preferencia por ciertos resultados sociales sobre otros, existen preferencias similares que parecen guiar a otros animales sin que ello implique necesariamente que a ellas subyazca el mismo sentido de obligación que guía a los seres humanos.

## REFERENCIAS BIBLIOGRÁFICAS

- Amici, F., Aureli, F. y Call, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Curr. Biol.*, 18, 1415-1419. <https://doi.org/10.1016/j.cub.2008.08.020>
- Aureli, F. y de Waal, F. B. M. (2000). *Natural conflict resolution*. Berkeley, CA: University of California Press. <https://doi.org/10.1525/9780520924932-003>
- Baier, A. C. (1991). *A progress of sentiments: reflections on Hume's treatise*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674020382>
- Barton, R. A. y Venditti, C. (2013). Human frontal lobes are not relatively large. *Proc. Natl. Acad. Sci. USA*, 110, 9001-9006. <https://doi.org/10.1073/pnas.1215723110>
- Bekoff, M. (2001). Social play behaviour cooperation, fairness, trust, and the evolution of morality. *J. Consc. Studies*, 8, 81-90.
- Beran, M. J., Savage-Rumbaugh, E. S., Pate, J. L. y Rumbaugh, D. M. (1999). Delay of gratification in chimpanzees (*Pan troglodytes*). *Dev. Psychobiol.*, 34, 119-127. [https://doi.org/10.1002/\(SICI\)1098-2302\(199903\)34:2<119::AID-DEV5><3.0.CO;2-P](https://doi.org/10.1002/(SICI)1098-2302(199903)34:2<119::AID-DEV5><3.0.CO;2-P)
- Bernstein, I. y Sharpe, L. (1966). Social roles in a rhesus monkey group. *Behaviour*, 26, 91-103. <https://doi.org/10.1163/156853966X00038>
- Black, M. (1970). *Margins of precision: essays in logic and language*. Ithaca, NY: Cornell University Press.
- Boehm, C. (1994). Pacifying interventions at Arnhem Zoo and Gombe. En R. W. Wrangham, W. C. McGrew, F. B. M. de Waal y P. G. Heltne (Eds.), *Chimpanzee cultures* (pp. 211-226). Cambridge, MA: Harvard University Press.
- Bräuer, J., Call, J. y Tomasello, M. (2006). Are apes really inequity averse? *Proc. Roy. Soc. Lond. B: Biol. Sci.*, 273, 3123-3128. <https://doi.org/10.1098/rspb.2006.3693>
- Brosnan, S. F. (2011). A hypothesis of the co-evolution of inequity and cooperation. *Front. Decis. Neurosci.*, 5, 43-55. <https://doi.org/10.3389/fnins.2011.00043>
- Brosnan, S. F. y de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, 425, 297-299. <https://doi.org/10.1038/nature01963>
- Brosnan, S. F. y de Waal, F. B. M. (2012). Fairness in animals: where to from here? *Soc. Just. Res.*, 25, 336-351. <https://doi.org/10.1007/s11211-012-0165-8>

- Brosnan, S. F., Schiff, H. C. y de Waal, F. B. M. (2005). Tolerance for inequity may increase with social closeness in chimpanzees. *Proc. Roy. Soc. Lond. B: Biol. Sci.*, 272, 253- 258. <https://doi.org/10.1098/rspb.2004.2947>
- Brosnan, S. F., Talbot, C., Ahlgren, M., Lambeth, S. P. y Schapiro, S. J. (2010). Mechanisms underlying responses to inequitable outcomes in chimpanzees, Pan troglodytes. *Anim. Behav.*, 79, 1229-1237. <https://doi.org/10.1016/j.anbehav.2010.02.019>
- Camerer, C. F. y Loewenstein, G. (2004). Behavioral economics: past, present, future. En C. F. Camerer, G. Loewenstein y M. Rabin (Eds.), *Advances in behavioral economics* (pp. 3-52). Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400829118-004>
- Churchland, P. S. (2011). *Braintrust: what neuroscience tells us about morality*. Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400838080>
- Coe, C. L. y Rosenblum, L. A. (1984). Male dominance in the bonnet macaque: a malleable relationship. En P. R. Barchas y S. P. Mendoza (Eds.), *Social cohesion: essays toward a sociophysiological perspective* (pp. 31-63). Westport, CT: Greenwood.
- Cords, M. y Thurnheer, S. (1993). Reconciliation with valuable partners by long-tailed macaques. *Ethology*, 93, 315-325. <https://doi.org/10.1111/j.1439-0310.1993.tb01212.x>
- de Waal, F. B. M. (1987). Tension regulation and nonreproductive functions of sex in captive bonobos (*Pan paniscus*). *Nat. Geogr. Res.*, 3, 318-335.
- de Waal, F. B. M. (1989). *Peacemaking among primates*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674033085>
- de Waal, F.B.M. (1992a). Appeasement, celebration, and food sharing in the two Pan species. En T. Nishida, W. C. McGrew, P. Marler y M. Pickford (Eds.), *Topics in Primatology, Vol. 1, Human origins* (pp. 37-50). Tokyo: University of Tokyo Press.
- de Waal, F. B. M. (1992b). Coalitions as part of reciprocal relations in the Arnhem chimpanzee colony. En A. Harcourt y F. B. M. de Waal (Eds.), *Coalitions and alliances in humans and other animals* (pp. 233-257). Oxford: Oxford University Press.
- de Waal, F. B. M. (1993). Co-development of dominance relations and affiliative bonds in rhesus monkeys. En M. E. Pereira y L. A. Fairbanks (Eds.), *Juvenile primates: life history, development, and behavior* (pp. 259-270). New York: Oxford University Press.

- de Waal, F. B. M. (1996). *Good natured: the origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674033177>
- de Waal, F. B. M. (1998 [orig. 1982]). *Chimpanzee politics: power and sex among apes*. Baltimore, MD: Johns Hopkins University Press.
- de Waal, F. B. M. (2000). Primates — A natural heritage of conflict resolution. *Science*, 289, 586-590. <https://doi.org/10.1126/science.289.5479.586>
- de Waal, F. B. M. (2006). The tower of morality: reply to commentaries. En S. Macedo y J. Ober (Eds.), *Primates & philosophers: how morality evolved* (pp. 161-181). Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400830336-011>
- de Waal, F. B. M. (2013). *The bonobo and the atheist: in search of humanism among the primates*. New York: Norton.
- de Waal, F. B. M. y Davis, J. M. (2003). Capuchin cognitive ecology: cooperation based on projected returns. *Neuropsychology*, 41, 221-228. [https://doi.org/10.1016/S0028-3932\(02\)00152-5](https://doi.org/10.1016/S0028-3932(02)00152-5)
- de Waal, F. B. M. y Ren, R. (1988). Comparison of the reconciliation behavior of stump-tail and rhesus macaques. *Ethology*, 78, 129-142. <https://doi.org/10.1111/j.1439-0310.1988.tb00224.x>
- de Waal, F. B. M. y van Roosmalen, A. (1979). Reconciliation and consolation among chimpanzees. *Behav. Ecol. Sociobiol.*, 5, 55-66. <https://doi.org/10.1007/BF00302695>
- Dindo, M. y de Waal, F. B. M. (2006). Partner effects on food consumption in brown capuchin monkeys. *Am. J. Primatol.*, 69, 1-9. <https://doi.org/10.1002/ajp.20362>
- Eberhard, W. G. (1972). The web of *Uloborus diversus* (Araneae: Uloboridae). *J. Zool. Lond.*, 166, 417-465. <https://doi.org/10.1111/j.1469-7998.1972.tb04968.x>
- Evans, T. A. y Beran, M. J. (2007). Chimpanzees use self-distraction to cope with impulsivity. *Biol. Lett.*, 3, 599-602. <https://doi.org/10.1098/rsbl.2007.0399>
- Fehr, E. y Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Q. J. Econ.*, 114, 817-868. <https://doi.org/10.1162/003355399556151>
- Flack, J. C., Jeannotte, L. A. y de Waal, F. B. M. (2004). Play signaling and the perception of social rules by juvenile chimpanzees. *J. Comp. Psychol.*, 118, 149-159. <https://doi.org/10.1037/0735-7036.118.2.149>
- Flack, J. C., Krakauer, D. C. y de Waal, F. B. M. (2005). Robustness mechanisms in primate societies: a perturbation study. *Proc. Roy. Soc. Lond. B: Biol. Sci.*, 272, 1091-1099. <https://doi.org/10.1098/rspb.2004.3019>

- Gallagher, J. (2004). Evolution? No: A Conversation with Dr. Ben Carson. *Adventist Review*, 26 (February). <https://www.adventistreview.org/archives/2004-1509/toc.html>
- Guth, W. (1995). On ultimatum bargaining experiments: a personal review. *J. Econ. Behav. Org.*, 27, 329-344. [https://doi.org/10.1016/0167-2681\(94\)00071-L](https://doi.org/10.1016/0167-2681(94)00071-L)
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. y McElreath, R. (2001). In search of Homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.*, 91, 73-78. <https://doi.org/10.1257/aer.91.2.73>
- Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.*, 3, 1-11. <https://doi.org/10.3389/neuro.09.031.2009>
- Herrmann, E., Keupp, S., Hare, B., Vaish, A. y Tomasello, M. (2013). Direct and indirect reputation formation in nonhuman great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*) and human children (*Homo sapiens*). *J. Comp. Psychol.*, 127, 63-75. <https://doi.org/10.1037/a0028929>
- Horner, V., Proctor, D., Bonnie, K.E., Whiten, A. y de Waal, F.B.M. (2010). Prestige affects cultural learning in chimpanzees. *PLoS-ONE*, 5, e10625. <https://doi.org/10.1371/journal.pone.0010625>
- Horowitz, A. (2009). Disambiguating the "guilty look": salient prompts to a familiar dog behaviour. *Behav. Process*, 81, 447-452. <https://doi.org/10.1016/j.beproc.2009.03.014>
- Hume, D. (1985 [orig. 1739]). *A treatise of human nature*. Harmondsworth: Penguin.
- Huxley, T.H. (1989 [orig. 1894]). *Evolution and ethics*. Press, Princeton, NJ: Princeton University.
- Jensen, K., Hare, B., Call, J. y Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proc. Roy. Soc. Lond. B: Biol. Sci.*, 273, 1013-1021. <https://doi.org/10.1098/rspb.2005.3417>
- Killen, M. y Rizzo, M. (2014). Morality, intentionality, and intergroup attitudes the origins of morality. *Behaviour*, 151, 337-359. <https://doi.org/10.1163/1568539X-00003132>
- Kitcher, P. (2006). Ethics and evolution: how to get here from there. En S. Macedo y J. Ober (Eds.), *Primates & Philosophers: How Morality Evolved* (pp. 120-139). Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400830336-009>
- Kitcher, P. (2014). Is a naturalized ethics possible? *Behaviour*, 151, 245-260. <https://doi.org/10.1163/1568539X-00003145>

- Koyama, N. F. y Dunbar, R. I. M. (1996). Anticipation of conflict by chimpanzees. *Primates*, 37, 79-86. <https://doi.org/10.1007/BF02382923>
- Kummer, H. (1995). *The quest of the sacred baboon*. Princeton, NJ: Princeton University Press.
- Logue, A. W. (1988). Research on self-control: an integrating framework. *Behav. Brain Sci.*, 11, 665-709. <https://doi.org/10.1017/S0140525X00053978>
- Mischel, W., Ebbesen, E. B. y Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *J. Person. Soc. Psychol.*, 21, 204-218. <https://doi.org/10.1037/h0032198>
- Osvath, M. y Osvath, H. (2008). Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: self-control and pre-experience in the face of future tool use. *Anim. Cogn.*, 11, 661-674. <https://doi.org/10.1007/s10071-008-0157-0>
- Palagi, E., Paoli, T. y Borgognini Tarli, S. (2006). Short-term benefits of play behavior and conflict prevention in *Pan paniscus*. *Int. J. Primatol.*, 27, 1257-1270. <https://doi.org/10.1007/s10764-006-9071-y>
- Price, S. A. y Brosnan, S. F. (2012). To each according to his need? Variability in the responses to inequity in non-human primates. *Soc. Justice Res.*, 25, 140-169. <https://doi.org/10.1007/s11211-012-0153-z>
- Proctor, D., Williamson, R. A., de Waal, F. B. M. y Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proc. Natl. Acad. Sci. USA*, 110, 2070-2075. <https://doi.org/10.1073/pnas.1220806110>
- Range, F., Leitner, K. y Virányi, Z. (2012). The influence of the relationship and motivation on inequity aversion in dogs. *Soc. Justice Res.*, 25, 170-194. <https://doi.org/10.1007/s11211-012-0155-x>
- Roma, P. G., Silberberg, A., Ruggiero, A. M., y Suomi, S. J. (2006). Capuchin monkeys, inequity aversion, and the frustration effect. *J. Comp. Psychol.*, 120, 67-73. <https://doi.org/10.1037/0735-7036.120.1.67>
- Russell, Y. I., Call, J. y Dunbar, R. I. M. (2008). Image scoring in great apes. *Behav. Process*, 78, 108-111. <https://doi.org/10.1016/j.beproc.2007.10.009>
- Singer, P. (1973). The trivality of the debate over "is-ought" and the definition of "moral". *Am. Philos. Q.*, 10, 51-56.
- Smith, A. (1937[orig. 1759]). *A theory of moral sentiments*. New York, NY: Modern Library.
- Subiaul, F., Vonk, J., Barth, J. y Okamoto-Barth, S. (2008). Chimpanzees learn the reputation of strangers by observation. *Anim. Cogn.*, 11, 611-623. <https://doi.org/10.1007/s10071-008-0151-6>

- van Hooff, J. A. R. A. M. (1972). A comparative approach to the phylogeny of laughter and smiling. En R. A. Hinde (Ed.), *Non-verbal communication* (pp. 209-241). Cambridge: Cambridge University Press.
- van Wolkenten, M., Brosnan, S. F. y de Waal, F. B. M. (2007). Inequity responses of monkeys modified by effort. *Proc. Natl. Acad. Sci. USA*, *104*, 18854-18859. <https://doi.org/10.1073/pnas.0707182104>
- Verbeek, P., Hartup, W. W. y Collins, W. C. (2000). Conflict management in children and ado-lescents. En F. Aureli y F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 34-53). Berkeley, CA: University of California Press. <https://doi.org/10.1525/9780520924932-006>
- Vollmer, P. J. (1977). Do mischievous dogs reveal their "guilt"? *Vet. Med. Small Anim. Clin.*, *72*, 1002-1005.
- von Rohr, C. R., Koski, S. E., Burkart, J. M., Caws, C., Fraser, O. N., Ziltener, A. y van Schaik, C. P. (2012). Impartial third-party interventions in captive chimpanzees: a reflection of community concern. *PLoS ONE*, *7*, e32494. <https://doi.org/10.1371/journal.pone.0032494>
- Westermarck, E. (1917[1908]). *The origin and development of the moral ideas*. Vol. 2. London: Macmillan.