Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), mayo-junio 2025,
Volumen 9, Número 3.

https://doi.org/10.37811/cl_rcm.v9i1

A DRIFT DIFFUSION MODEL APPROACH TO MORAL DECISION-MAKING: TOWARD A COMPUTATIONAL FRAMEWORK FOR ETHICAL DILEMMAS

UN ENFOQUE DE MODELO DE DIFUSIÓN DE DERIVA PARA LA TOMA DE DECISIONES MORALES: HACIA UN MARCO COMPUTACIONAL PARA DILEMAS ÉTICOS

Carlos Ledesma – Alonso
Student, Faculty of Bioethics, Universidad Anáhuac Campus Norte, Estado de México

DOI: https://doi.org/10.37811/cl_rcm.v9i3.18524

A Drift Diffusion Model Approach to Moral Decision-Making: Toward a Computational Framework for Ethical Dilemmas

Carlos Ledesma – Alonso[1]

carlos_ledesma@anahuac.mx

https://orcid.org/0009-0005-3744-6422

Student, Faculty of Bioethics, Universidad Anáhuac Campus Norte, Estado de México

Huixquilucan, , Mexico

ABSTRACT

Can ethical dilemmas be resolved through computational models? In this paper I propose a novel integration of the Drift Diffusion Model (DDM) with a value-based framework for moral reasoning. Based on insights from neuroethics and cognitive neuroscience, I argue that moral decisions — especially under conditions of uncertainty, irreversibility, and emotional interference— can be modeled as processes of noisy evidence accumulation. I introduce the concept of an *ethical balance* mainly composed of three evaluative variables: sentience, intentionality, and innocence. These inputs are mapped to DDM parameters such as drift rate, decision threshold, and starting point bias. Through illustrative moral scenarios, I show how this framework can both predict and simulate moral judgments. The goal is not to replace normative ethics, but to demonstrate how scientific modeling can enhance our understanding of how moral reasoning unfolds in the brain. This maybe opens the path to a computational ethics grounded in real cognitive processes.

Keywords: drift diffusion model, neuroethics, decision – making, ethical balance, moral dilemmas

[1] Autor principal
Correspondencia: carlos_ledesma@anahuac.mx

Un Enfoque de Modelo de Difusión de Deriva para la Toma de Decisiones Morales: Hacia un Marco Computacional para Dilemas Éticos

RESUMEN

¿Se pueden resolver los dilemas éticos a través de modelos computacionales? En este artículo propongo una integración novedosa del Modelo de Difusión de la Deriva (DDM) con un marco basado en valores para el razonamiento moral. Con base en los conocimientos de la neuroética y la neurociencia cognitiva, sostengo que las decisiones morales —especialmente en condiciones de incertidumbre, irreversibilidad e interferencia emocional— pueden modelarse como procesos de acumulación de evidencia ruidosa. Introduzco el concepto de un *equilibrio ético* compuesto principalmente por tres variables evaluativas: sensibilidad, intencionalidad e inocencia. Estas entradas se asignan a los parámetros de DDM, como la velocidad de deriva, el umbral de decisión y el sesgo del punto de partida. A través de escenarios morales ilustrativos, muestro cómo este marco puede predecir y simular juicios morales. El objetivo no es reemplazar la ética normativa, sino demostrar cómo el modelado científico puede mejorar nuestra comprensión de cómo se despliega el razonamiento moral en el cerebro. Esto tal vez abra el camino a una ética computacional basada en procesos cognitivos reales.

Palabras clave: modelo de difusión de deriva, neuroética, toma de decisiones, balanza ética, dilemas morales.

Artículo recibido 11 junio 2025
Aceptado para publicación: 30 junio 2025

INTRODUCTION

Science has explained most natural phenomena — from planetary motion to subatomic interactions — with remarkable precision. From Newton's laws and Maxwell's equations, to Einstein's relativity and Dirac's quantum predictions, physics offers deep explanatory power. Other fields such as medicine and biology have also solved foundational questions, including the origin of species through Darwinian evolution. This broad explanatory capacity suggests that science may also illuminate the nature of morality.

In this article, I analyze and discuss how the scope of science reaches areas where it was thought that only ethics and morality —both philosophical currents— had interference. In addition, I argue that, by understanding the neurobiological mechanisms that underlie phenomena such as altruism or compassion, both components of an ethical dilemma, neuroethics (a relatively recent emerging science) can help us unveil the phenomenon behind morality and even converge in decision-making with ethical implications.

Concept of factual science and its historical context

To understand how ethics should be added as just another science, we must have a complete definition of science. To consider ethics as a science, we must adopt a rigorous definition. According to Mario Bunge (2013), a factual science is defined by a real domain, a community of researchers, a realistic epistemology, and verifiable methods and theories. Ethics meets these conditions when grounded in cognitive neuroscience, thereby justifying neuroethics as a scientific discipline (Bunge, 2013).

Science, contextualized from a historical perspective, according to Thomas Kuhn, goes through different transitions: pre-scientific, normal, and revolutionary. The transition from one transition to another is carried out by the change of scientific paradigms and in this way, science advances (Kuhn, 1994) . However, Kuhn's thesis and his paradigms do not seem to be entirely clear mainly due to the fact what that Kun's consider revolutions are just great discoveries not revolutions by itself.

Despite the great scientific advances, from Archimedes to Einstein, epistemologically speaking it has been affirmed that science has its limits, or at least that is what a few philosophers of science argue.

The most recognized is probably the philosopher Paul K. Feyerabend. In one of his works, he stated that Galileo Galilei could not have made the lunar observations and drawn his conclusions, if he had carried out the steps of the scientific method as such (Feyerabend, 1987). In addition, Feyerabend erroneously asserts that some pseudosciences and pseudotechniques such as homeopathy, psychoanalysis and astrology, should be considered sciences, under the assumption that the scientific method does not exist as such, calling it a gnoseological anarchism or *anything goes*. It is clear that Feyerabend's vision was compatible with pseudoscience, making his thesis redundant, unclear and irresponsible.

Despite Feyerabend's failed attempts to categorize pseudosciences and pseudotechnologies as sciences, science has always been successful in reaching a result that works, even if it is not carried out experimentally. An example of this was the achievement of physicist Peter Higgs who theoretically hypothesized a boson in the 1960s in a mathematical way, and it was confirmed experimentally in 2012 by the powerful hadron collider in Bern, Switzerland. In addition, the objective of science is the continuous improvement of its main products (theories) and means (techniques), as well as the subjection of ever-increasing territories to their power (Bunge, 2000).

In this regard, Bunge himself argues that there are two types of sciences: formal and factual (Bunge, 2018). These are formal sciences, such as logic or mathematics, which provide us with verifiable and complete information about reality, but do not deal with facts, and factual science, where their statements need to be verifiable through experience.

Adjusting our central argument to Bunge's concept, the formal sciences; demonstrate, and the factual; confirm. While the demonstration is complete, the verification is incomplete (*On. Cit.* Bunge M, 2005). In this sense, I emphasize again that science can satisfactorily explain the phenomenon of ethics.

Digression: A poorly posed question gives the wrong answer

One of the most inadequate questions is: Can science resolve moral dilemmas? Does not have an answer of a dichotomous nature, since, if so, the answer would be yes and no, which is unsatisfactory. Consequently, it is a poorly posed question that has implications that do not hold water. The issue is not that it is more complex, it is just different. To ask whether science can solve ethical dilemmas would be to reduce science to its method. To put this in perspective, let's look at it this way: mathematics is to physics what the scientific method—or its variants—is to science. But mathematics is not physics; just as the scientific method is not science. That said, to answer the question, it must first be posed properly.

Given that science, in particular neuroscience and its branches, can study phenomena such as empathy, reason, guilt and decision-making, then it can be deduced that it helps to unveil the ethical phenomenon, that is, from a purely scientific perspective. Considering that empathy, reason, guilt and decision-making are neurophysiological processes that are used to solve ethical dilemmas, we can argue with a certain degree of certainty that science can deal with the explanation of the origin and neurobiological mechanisms that underlie morality, but it is not in charge of acting as a judge and ruling what should be done in this or that situation.

It should be added that if we wish to categorize ethics as a scientific phenomenon, it must materialize, therefore dualism must be replaced. The psychoneural dualist view is incompatible with the ontological reduction that asserts that all that is mental is the same as the neural. Let us remember the most forceful thing that Bunge tells us about dualism (Bunge, 2010). First, dualism is *conceptually imprecise*, that is, the very concept of action is only defined by material things, hence, to say that mind-body interaction is untenable, because being immaterial mind is insensitive to physical stimuli. Second, dualism is *experimentally irrefutable*, since something immaterial cannot be manipulated and, consequently, experimented with. Third, dualism is *incompatible with cognitive ethology*, because the solid evidence comes from anatomical and physiological studies and whose relationship brings us closer to nonhuman animals, especially primates, and does not distance us from them. Fourth, dualism *violates a fundamental physical law*, namely the law of conservation of energy. Fifth, it *confuses scientists*, since when talking about "correlates that underlie x phenomenon" it is as much as saying that the legs *implement* walking or that the stomach *instantiates* digestion, remember that there is no function without an organ and no organ without functions. In other words, psychoneural dualism is scientifically and philosophically untenable.

With this it is more than clear that the faculties of the brain such as thoughts, learning, morality or free will are highly complex biophysical processes that are explained by studying the emergent properties of neurons and other brain components (glia, neurotransmitters, etc.), their agglomerations, their functions and their interactions through the sum of one or more scientific disciplines and not through the invocation of immaterial entities.

A proposal for a mathematical model as a tool to *elucidate* ethical problems

A pitfall-free and more appropriate question would be whether science can explain and substantiate ethical phenomena, not solve them. Ethical problems are not solved as in the formal sciences (e.g., mathematics), but they are reasoned and deduced based on accumulated evidence. The answer to the question is yes, from a neuroscientific perspective, as mentioned above. Jumping from this explanation to the question of whether science can solve ethical problems would be like trying to explain consciousness, only by studying the anatomical-functional unit that generates it, the neuron. The phenomenon is much broader, since it is emerging. However, with a certain process that involves reasoning, we could try *to elucidate* and thus try to conclude a simple ethical conflict – in the medium to long term – using a mathematical model.

For this, I imagined a very simple panorama, consisting of four components, in order to subtract the fundamental parts that make up an ethical problem. Let's imagine a scenario where there are four participants: a pig, an apple tree and two human beings. There are only a couple of rules, and they are as follows:

1. Each individual can only eat one of the others to survive.
2. The most ethical scenario will be the one that has generated the least damage.

Scenario I

One human kills the pig, the other eats the apple tree, consequently, both humans are left alive.

Scenario II

One human kills the other human to feed the pig, then he eats the apple tree. The pig and the human are still alive.

Scenario III

One human wish to kill the pig, the other human intercepts and kills it. The pig eats the dead human and the remaining human the apple tree. The pig and the human are still alive.

Scenario IV

One human gives life to feed the pig, the other human eats the apple tree. The pig and the human are still alive.

Scenario V

The two humans are slaughtered, and the pig gets to choose who to eat. Only the pig remains alive.

From the above panorama, different variables can be subtracted to consider and dilute an ethical dilemma. One of them is *sentience*. By definition, an organism is sentient if it has a nervous system complex enough to perceive pain and process it and, consequently, generate suffering. The next is *innocence*. By definition, it is considered an innocent being, if it does not seek to harm any sentient being without any logical and convincing reason – innocent or not – or the damage it generates is minimal and unintentional. The last is *intentionality*. A being is intentional, if he has the cognitive capacity to make decisions, whether good, bad or neutral.

Therefore, this thought experiment reveals key variables for moral evaluation: sentience, innocence, and intentionality. Assigning binary values to each (1 or 0), we can derive an "ethical valence" score. In Scenario III, a human defends a sentient, innocent being, resulting in two high-valence survivors. Scenario IV is similarly ethical due to altruistic sacrifice. This algorithmic structure lays the foundation for the proposed ethical balance.

For example, a pig is a sentient, innocent and intentionally good being, it has a valence of 3, because it does not harm. The apple tree has 1 point (innocence), because although they sense (perceive) and have intention (survive), they are in zero degrees compared to a mammal. In humans, their valence can range from having 1 point to having 3 points. In scenario III, the human intends to defend a being of valence of 3 (innocent, sentient and well-intentioned), eliminating one that has bad intentions. Thus, in panorama III, two beings of valence of 3 are still alive. However, scenario IV may be just as good as scenario III, because a human sacrificed himself to feed a sentient, innocent, well-meaning being. I extend this better in the next section.

This panorama can be extended in complexity, if we consider a fourth factor: *irreversibility*. By definition, an event is irreversible if none of its components can be restored in a primordial way, e.g. death. Which means that we can confer gradation on any ethical conflict in any area of life. If we want to use this ethical scale in any ethical or moral landscape, irreversibility must be considered in order to confer the level of importance on it. In the hypothetical case of the pig, humans and the apple tree, there will always be death, something irreversible. This means that, in any ethical or moral conflict, there will always be *damage* (last factor), if there is an irreversible factor in the background, such as death. The other factor to add importance to an ethical conflict will be the damage it causes, and not the happiness it causes. Hence, this method should not be considered strictly utilitarian. If there is no harm, there will be well-being, which is not the same as happiness.

Towards ethical computation: Simulation Framework for the Ethical Drift Diffusion Model

Once the variables have been extracted, I propose the use of the mathematical model called *the Drift Diffusion Model* (DDM). DDM is defined by a series of mathematical equations that contain various parameters to which different values can be assigned, in turn, each of these parameters can be adjusted in the model, which affects its subsequent behavior (Myers et al., 2022). For example, the higher the value of the drift rate (μ), the faster the decision will be made. On the other hand, decision boundaries (represented ±θ) determine the amount of evidence accumulated to make the decision. In this sense, it has been studied the influence of the affect – as – information may determine the drift rate along with other choice-relevant attributes (Roberts & Hutcherson, 2019). Also, this model has been used in patients with lesions in the orbitofrontal cortex in a decision-making task (Peters & D'Esposito, 2020).

To advance the empirical applicability of the proposed *ethical balance model*, I´ve to summaries a simulation-based implementation using the Drift Diffusion Model (DDM). The simulation aims to model moral decision-making under conflict by quantitatively mapping ethical variables to DDM parameters.

Each option in a moral dilemma is evaluated according to five key ethical variables:

Table 1. Ethical Variables as Inputs. A weighted ethical value (V) is calculated for each actor or action as: $v = \alpha S + \beta I + \gamma N - \delta R - \epsilon D$; where *α, β, γ, δ, ε* are tunable weights reflecting ethical emphasis. The values in S, R and D oscillate between 0 to 1.0 (e. g. 0.8), this is mainly based on the context of the dilemma because it's a conceptual estimation.

Variable	Symbol	Type	Scale	Description
Sentience	S	Quantitative	[0.0 – 1.0]	Capacity to feel pain/suffering
Intentionality	I	Binary	0 or 1	Whether harm was intentional
Innocence	N	Binary	0 or 1	Whether the agent is morally blameless
Irreversibility	R	Quantitative	[0.0 – 1.0]	Degree of irreversibility of the consequence
Damage	D	Quantitative	[0.0 – 1.0]	Magnitude of harm inflicted

I have to make emphasis that the DDM decision process compares two options (A and B), with the following parameter mappings:

Table 2. Mapping to DDM Parameters.

DDM Parameter	Source	Meaning
Drift rate (μ)	$VA - VB$	Bias toward the ethically superior option
Decision threshold (θ)	Function of max (R, D)	Stricter criteria under high-stakes decisions
Starting point (z)	Prior bias, e.g., preference for humans	Encodes pre-existing moral inclinations

Let us now examine an illustrative and very simple application with steps of the ethical balance with DDM.

1. Found a dilemma.

A: Save an animal who is in danger of die

B: Save a serial killer in the same conditions

2. Assume the following input values according with step 1:

Option	S	I	N	R	D
A	1	1	1	0.8	0.1
B	1	0	0	0.8	0.3

Using weights: $\alpha = 1.0$, $\beta = 1.0$, $\gamma = 1.0$, $\delta = 1.5$, $\varepsilon = 1.0$

3. Use the ethical balance based in DDM.

VA = 1 + 1 + 1 - 1.5(0.8) - 0.1 = 1.7

VB = 1 + 0 + 0 - 1.5(0.8) - 0.3 = -0.5

μ = VA - VB = 2.2

4. Interpretate the results:

Conclusion. The high drift rate ($\mu = 2.2$) favors option A, suggesting a fast and confident decision. If μ is near 0, the moral dilemma would be interpreted as difficult or ambiguous, nevertheless, if μ have high positive or negative numbers, the resolution of the dilemma would be easier.

This simulation can be implemented using standard DDM tools in Python (e.g., PyDDM), enabling analysis of a) response times, b) accuracy/conflict and c) bias under varied initial values. Such a simulation could be extended to include family ties, temporal pressure, or speciesism by adjusting initial bias (z) or noise. This scenario maybe could be considered simple and improbable, but the core of this model is to design a very precise and clear context. Although the ethical scenarios that occur at any given moment – in many areas, both natural and artificial – vary in complexity, eventuality and temporality, they have practically the same components and must be dimensioned in degree of importance. Consequently, I propose that all ethical dilemmas should be categorized in degree of importance and that this is a function of *irreversibility,* and the damage generated that can presumably be calculated, as in the example above. Note: the difference between valences (V) is not decided by us, but rather by the decision-making process: whether it will be easy or difficult, fast or slow, confident or difficult.

Empirical and Simulated Approaches to Moral Modeling

One of the most viable ways to explore the scope of the Ethical Balance based on DDM is through the design of behavioral experiments involving hypothetical moral dilemmas. These scenarios can be carefully constructed to incorporate varying degrees of sentience, intentionality, innocence, irreversibility, and damage—dimensions that, according to the ethical balance proposed here, are fundamental in shaping moral decisions. Participants would face dilemmas where these variables are subtly embedded, and their decisions could then be contrasted with the predictions generated by the model.

At the same time, the model lends itself to theoretical exploration through simulations. By assigning plausible values to each ethical input and adjusting their respective weights, one can anticipate the direction and speed of moral choices. As expected from DDM principles, a high difference in ethical valence between two options results in faster and more confident decisions, while scenarios with similar values generate hesitation and cognitive conflict.

The purpose is not to emulate moral philosophy, but to better understand how real cognitive processes might operate when we are forced to act. By comparing simulated patterns with empirical responses, the ethical balance can serve as a useful and testable framework for studying how the brain—and perhaps even artificial agents—approach moral conflict. *Review its simulation in appendix A*.

Limitations of the ethical balance based on DDM

The ethical balance has its scope, just as the scientific method is not irrefutable, nor complete, but it brings us closer to the factual truth. It cannot be used satisfactorily if we consider other weighty factors such as feelings. Whether due to family ties, friendships, genetics or determined by strong beliefs, feelings put such noise in the balance that in any situation we will always prefer to save our loved one, ignoring any normative ethics.

In nature we find clear examples of this. A hyena in the African savannah would mindlessly devour a lioness protecting a pride of young. The hyena does not reason or perceive empathy for unprotected offspring that are likely to die in starving or painful conditions without the help of their mother. But we must not fall into the trap that, in nature, every act is good. This would lead to the naturalistic fallacy proposed by G. E. Moore.

Darwin's theory of natural selection sometimes defies the utilitarian precept by putting the life of a single offspring above other species. According to Richard Dawkins, genes are entities with selfish motivations (by wanting to replicate themselves at all costs) and this explains certain traits in animal behavior, such as sexual reproduction or altruism (Dawkins, 1989). However, this concept turns out to be erroneous, since genes are primarily chemical structures that also depend on the environment (epigenetics).

Another scenario where the ethical balance will fail is if we consider a very short time as a factor. The time you spend deciding whether or not to unplug a patient in a coma is not the same as attacking and killing a thief who breaks into your house. And that is what I mean by ethics with temporal gradation, namely, short, medium and long term.

In fact, considering Daniel Kahneman's two decision-making systems, one fast and intuitive and the other slow and logical (Krämer, 2014), I assert that, in order to solve an ethical dilemma, we must oscillate depending on how we make the decision, considering factors such as: time, irreversibility and damage. In fast-paced situations, we act according to system 1, but in a situation that takes longer, we can apply logical and systematic reasoning (system 2).

From this we can conclude that, in system 1 we use regions of the brain such as the limbic system; largely responsible for generating emotions and feelings, and system 2; which uses regions such as the prefrontal cortex, the seat of executive thinking and logical reasoning.

In conclusion, the decision making of an ethical dilemma depends on its temporal gradation, namely, short, medium and long-term resolution. In addition, the resolution of moral problems by means of the *ethical scale* or the ethical balance based on DDM does not pretend to be complete and true, only to approximate the truth.

When the Ethical Balance remains at zero: Justice Beyond the Dilemma

Justice, whether legal or moral, is often arbitrary and insufficient to proportionally address harm. In ethical dilemmas, some degree of damage is inevitable, and ideally, compensation should be proportional. When the ethical balance remains neutral—as in accidental dilemmas—justice must address irreparable loss. Recognizing the shared ontological origin of living beings challenges human exceptionalism and reinforces the role of science in guiding fair responses.

In 2016, a terrible accident occurred. At a zoo located in Cincinnati, United States, a 3-year-old boy accidentally fell into a gorilla enclosure. After about 10 minutes, a burst of bullets struck the gorilla, killing him instantly, saving the boy's life. What did this act mean? Well, the police considered that the life of the minor was worth more than that of the gorilla. This was not to be evaluated in this way, because the gorilla had no intention of hurting the child. Two sentient beings, innocent and not ill-intentioned, end up in a difficult situation. The ethical balance is not moving anywhere, what next? After the decision is made, justice must be done. What kind of justice should there be if a life cannot be recovered? Hence the premise that we could adopt is that human ontology is neither superior nor essentially different from that of practically any living being. In fact, it is generally accepted that the origin of every living organism lies in a common ancestor called LUCA (*Last Universal Common Ancestor*) whose metabolic reconstruction includes common pathways throughout the tree of life (Goldman & Becerra, 2024). As Darwin stated, [humans] are different in degree but not in class.

Justice must consider the above factors, namely sentience, innocence, and intentionality, and from these, the harm caused, and the costs involved in claiming the harm. At the same time, I propose that justice should rely heavily on science and its methods.

The Neurobiological Foundations of Moral Cognition

In the past, science was considered to be solely descriptive and ethics normative. The truth is that ethics can take on a scientific nuance. Ethics alone can be considered as a semi-science, however, by adding several fields of research such as neuroscience, physiological psychology, sociology and evolutionary psychology, neuroethics emerges and should be considered as an emerging science.

The most famous case in neuroscience that gave rise to hypothesizes that our morality originates from how our neurons are connected was the case of Phineas Gage, a miner who had an accident in 1858. As a result of an explosion, an iron bar embedded itself from his jaw and pierced his frontal lobe, causing irreversible damage. Colleagues and friends attest that before the accident, Gage was kind, hardworking, and a good friend; however, after the event, Gage became bitter, angry, and lazy. It was clear that something had happened to his brain.

A recent study showed that two adult patients with the *Gage syndrome* who had severe lesions in the prefrontal cortex, caused before 16 months of age, had impaired social behaviors, insensitivity to consequences, and suffered from defective moral reasoning (Anderson et al., 1999). Another study carried out by Michael Gazzaniga and his research group showed that neuroimaging studies indicate that there are lateral brain mechanisms in the temporoparietal junction that are involved in complex moral and social reasoning and that subjects with a previous callosotomy (surgical removal of the corpus callosum) showed a clear alteration in their moral judgments (Miller et al., 2010).

While common sense often guides moral judgments, neuroimaging studies show that ethical reasoning activates brain regions like the medial temporal, cingulate, and frontal gyri. These findings suggest that moral cognition involves more than intuition—it requires theory of mind and higher-order processing distributed across cortical and subcortical networks. (Garrigan et al., 2016).

In addition, another experiment that used functional magnetic resonance imaging, analyzed the brains of participants while they solved personal moral dilemmas and showed that brain areas such as the prefrontal cortex and the anterior cingulate cortex were activated; both related to abstract reasoning and cognitive control (Greene et al., 2004). Common sense is rather intuitive, and many times moral reasoning requires little more than pure intuition. In fact, it has been shown that there are cases where a person uses theory of mind cognition to make a moral judgment (Knobe, 2005; Moll & de Oliveira-Souza, 2007; Moretto et al., 2010), for this, several cortical and subcortical regions must be used.

Other natural causes that show that our morality depends to a large extent on our proper neuronal functioning are neurodegenerative diseases (Alzheimer's or Parkinson's disease). It has been shown that people suffering from frontotemporal dementia manifest violations of moral rules and norms, a fact that may be due to what researchers call *moral agnosia*, or an inability to recognize right from wrong (Mendez et al., 2005). In this regard, it has been proven that patients with Parkinson's disease have alterations in moral and social aspects, that is, not only the motor aspect (Santens et al., 2018).

Likewise, neuroscience —and its branches— can satisfactorily explain the origin of the moral sense that we animals perceive, which are mere manifestations of evolutionary impulses such as doing no harm, justice or equity, community, authority and purity (Mendez, 2009). In this way, we can argue that the origin of a basic moral sense arose in an evolutionary way and is shaped by the interaction between species (same or different), the environment and self-interests (physiological and psychological).

The affective and emotional components of the brain allow us to understand the situation of another subject (e.g., empathy). Such behaviors have been demonstrated in rats, corvids, monkeys, elephants, and human babies. Prosocial mechanisms have their evolutionary roots, and many social norms that are based on them help the correct and adequate interaction between species (Wu & Hong, 2022).

Consequently, it can be said that an animal is endowed with an *intentional* brain, because they have both positive (e.g. cooperating) and negative (e.g. killing for fun) volitional actions or purposes. In fact, it has been shown that, in primates, such as chimpanzees, qualities such as empathy, fairness and loyalty in the group have been described (Chung et al., 2021; de Waal, 2021), all components that lead to good morals. And if analyzed in a sociocultural way, in all regions of the world, cultures and societies have an almost universal basic distinction between what is good and bad. In other words, social, cultural, intellectual, political, and biological changes have shaped and, in some way, (e.g., creation of civilizations or secular education) the human moral sense, which is rather moral *reasoning,* although this does not occur in all human actions (e.g., speciesism or climate change). In fact, according to the theory of moral realism, it demands an objective distinction between right and wrong, good and bad, based on reasoned arguments explained in the work of D. Brink (Brink, 1989).

On the other hand, a study that worked with Pavlovian-type learning mechanisms, based on the role played by the amygdala in decision-making, showed that this can be useful, namely, invoking inherent prosocial tendencies (e.g., cognitive and emotional empathy and various forms of altruism), as well as prioritizing reciprocity between individuals and fostering cooperation (Seymour & Dolan, 2008). This shows that the results of science, in particular, behavioral sciences, can help to promote prosocial behaviors. In other words, the results that science reveals can be used in favor of moral reasoning.

Finally, as has been proven sufficiently, animals are equipped to solve a myriad of problems, even facing our own nature. But you can also be good or ethical, learning to be so. A clear example of this is the person who has been eating meat all his life and suddenly discovers solid arguments that force him to stop consuming animals. That is, ideologies can determine a person's moral orientation. It is for this reason that we must equip our brain with skepticism and good judgment in order to make the best decisions.

Factors that distort ethical judgment

In an impartial environment (e.g., hospitals or clinics), when making a decision that has ethical implications, there are times when you must be as *clean* as possible, that is, there are no family ties or friends involved. In the professional field, feelings and emotions can manipulate our decisions with what seem like compelling reasons, but in reality, they are not. An example of this would be the doctor who falls in love with one of his patients, and, consequently, will give him more importance and thus preference than other patients. This is what I mean by saying as clean as possible.

However, others claim that the availability heuristic, and other types of biases, can be useful in medical practice, as long as factors such as experience or personality are taken into account (Whelehan et al., 2020). In turn, Marewski & Gigerenzer (Marewski & Gigerenzer, 2012) argue that heuristics are more useful in medicine than is commonly believed, as they save a lot of time and are accurate. However, in this sense, I assert that, although heuristics can work in a certain percentage, a set of powerful algorithms such as *deep learning* using convolutional neural networks have incredible accuracy in diagnosing diseases, even when compared to medical experts in the area (Richens et al., 2020). Even the ability of *deep learning* to analyze images can provide support for physicians decisions and improve the accuracy and efficiency of various diagnostic and treatment processes (Chan, Hadjiiski, et al., 2020; Chan, Samala, et al., 2020). This emphasizes the fact that the less biased we can obtain better results in most cases. These experiments support, at least indirectly, the use of machines for the resolution of some moral dilemmas.

The beliefs that the brain generates as a result of external and internal stimuli, experiences, and their subsequent processing, can be as labile and fragile as they are solid and immovable. This happens because the brain is susceptible to the information presented to it. The belief that the earth revolves around the sun is verifiable, not only through experience, but also through mathematics. However, the belief in fictitious entities such as ghosts is irrational, illogical and unfounded, despite the fact that our brain contains experiences that affirm it. This mental phenomenon is known as *Cognitive bias*, which is nothing more than an erroneous interpretation of reality, derived from failure or an absence of a systematic, rational and verifiable process. Conclusions drawn from cognitive bias give rise to fallacies and can lead to poor decisions.

However, there are studies in which, using Bayesian-type probabilistic models, the results show that they contribute to an empirical explanation of how human inductive learning biases can influence cultural transmission (Thompson & Griffiths, 2021). In other words, humans rely on the learning of their groupmates rather than learning them by environmental feedback. So, it can be said that some biases are less harmful than others.

An important factor to consider in human decisions is noise. According to Kahneman, noise in a system is an undesired variability in a judgment (Kahneman et al., 2021). Kahneman affirms that noise is present in every decision we make, however; There are strategies that can help us reduce noise. While bias results in a consistent but erroneous conclusion, noise confers inconsistency in our judgment, so our conclusion is likely to be wrong as well.

Emotional state is one of the factors that influence the most when we make a decision. There are emotions even when we think rationally, because cognitive-emotional interactions have a high degree of connectivity, so they cannot be separated (Pessoa, 2008). It has been demonstrated, in both non-human and human animals, that cognition-emotion links play an important role in the generation of emotional states and these, in turn, influence cognition, inducing different types of biases (Paul et al., 2005). In fact, one hypothesis that supports the high degree of connectivity between structures related to cognition and emotion is the somatic marker hypothesis. Damasio (Damasio, 1996) argues that somatic markers arise in bioregulatory processes, emotions and feelings and all this is related to the body-state structure, therefore, the somatic marker hypothesis by definition does not accept that decision-making is purely rational or cognitive.

DISCUSSION

The ethical balance based on the mathematical model DDM can be modulated through the context, and the variables involved. It's limitations are determined by the complexity and what the human consider damage and innocence. In a professional environment like a hospital, perhaps, it can be used, but I must highlight that this model does not replace human decisions, only can help to secure a response.

On the other hand, science and it's product (knowledge) can help us clarify the resolution of an ethical dilemma, because it can explain the neurobiological mechanisms that underlie events to do good, such as compassion and empathy, or doing evil, such as anger and revenge, and because with their findings, we can even become aware of how we should treat our world in general. In addition, the positions defended in this work can be adopted or debated, since they are proposals argued with evidence but not assertions made by simple authority.

Neuroscience satisfactorily explains how emotions, feelings, beliefs, and personal interests can generate biases that shape decision-making and cloud proper ethical judgment, so that we cannot always rely on simple common sense to resolve moral dilemmas. Common sense is essentially intuitive, so a dilemma that requires more than one way of perceiving a problem is unsatisfactory. By way of digression, in a situation where a quick decision must be made, we cannot use any formal methodology, namely slow methodology; only to act by intuition quickly, that is; use Daniel Kahneman's system I.

According to Damasio's hypothesis (Bechara et al., 1999), even if a person tries to be completely rational, there will always be a degree of emotion or feeling involved. But in practice, a person is able to affirm that he can make a rational decision, by following a concatenated series of reasoning. Perhaps the degree or percentage of emotion is what can determine the direction of a decision.

Neuroscientist Sam Harris has made a good case regarding the involvement of science around human values. He came very close to the question by asserting that the increase in the well-being of living beings is explained from a scientific perspective (Harris, 2011) and with this science can explain human values. However, there are philosophers who are detractors of this position.

Brian Earp asserts that a crucial question is how science defines well-being. In addition, he affirms that what Harris calls science is in reality philosophy (Earp, 2015). In my book *My Morals Groping* (Anonymous, 2022), I tried to refute the *crucial* question to which this philosopher refers, about how science can define well-being. Neuroscience not only defines well-being, also explains it and even demonstrates it. Although it is neurophysiologically difficult to define well-being, it can be evaluated qualitatively. Even so, there are neural traits that can define a state of well-being.

In this article, I used a kind of algorithm based on various factors, but always considering evidence that science has helped to clarify. More in favor of beings who cannot express in words what they feel: animals. Such is the case of how the brain of animals is more sensitive and cognitive than we used to believe.

In this regard, a pilot study investigated the empathy that people perceive towards animals and determined that a more empathetic attitude is linked to knowledge of skills, cognition and sensitivity that people have about animals (Cornish et al., 2018). Heyes (Heyes, 2018) argues that research conducted on human animals and infants suggests that the mechanism of empathy and emotional contagion is built through social interaction, and that, in turn, it is as versatile as it is fragile. As a result of this, he asserted that the more we know about the brain of animals, the more ethical our attitude towards them is. In this way, it can be concluded that the advance of science can have the effect of a substantial improvement in our respect for living beings and thus make us more ethical. Thus, science does have to do with ethical issues.

In addition, logical reasoning does not have a patent. It is not typical of philosophy, nor of science, although both use it to reach conclusions or theories. But, unlike philosophy, science uses a methodology to reach a result. In this sense, and only in this, philosophy serves as a tool for science.

The use of methodologies to solve bioethical dilemmas is not new (De los Ríos-Uriarte, 2017). There are several methods that serve as a basis for the deliberation of a problem in clinical medicine, such as the Thomasma-Pellegrino method, or the Jonsen-Siegler four-topic method (Jonsen, 1982). Both are to some extent useful and frequently used; At the same time, they consider values, principles, and medical findings, but organizing information does not necessarily mean making a decision. At some point in both methodologies, deliberation has to be deliberated, and the inclusion of values certainly introduces noise to decision-making. Instead, the use of the mathematical drift diffusion model goes to the heart of the matter: decision-making in the brain based on evidence and certain variables.

It should be added that normative ethics can guide us to be better people with our environment, as well as with biotic and abiotic systems, and not only with other people. That is why there are rules and norms that set limits on certain behaviors, that is, in the legal framework. Even so, there are actions that will never be penalized despite the fact that their nature is objectively bad. In this regard, the philosopher and ethicist Peter Singer stated that, if an individual has the moral obligation to do something, he should do it, otherwise, it can be considered immoral (Singer, 1972). While it may cause some discomfort to admit that one is immoral because one has not planted trees, donated to a charity, or adopted an animal in need, one is responsible on a small but perceptible scale. Consequently, that someone can be considered immoral.

From this I can also suggest that, if a person has physical, intellectual, economic, or academic abilities above the statistical average, he has a moral obligation to do something known as effective altruism, and not only with other humans, but with all living things. A privilege obtained by fortuitous events or generated by the combination of personal success and luck, should be shared (not with everyone) and not limited (based on one's own well-being). This is without defocusing personal interests either, as long as they do not border on banality.

On the other hand, the use of this type of algorithms and mathematical models such as the drift diffusion model has good possibilities to help us solve some moral dilemmas, however, the technical difficulty will result in the assignment of values to the parameters, as well as their calibration according to the context. With the use of new computer technologies and more advanced algorithms such as *Deep Learning* and its neural networks that self-update to learn on their own, the idea that they can help humans solve problems that were intended only for the human brain is not as far-fetched as it seems.

Appendix A. Technical Simulation of the Ethical Balance

To demonstrate the internal coherence and potential applicability of the ethical balance based on Drift Diffusion Model (DDM), a basic simulation was conducted using three moral dilemmas constructed with hypothetical yet plausible values. Each ethical dilemma was structured around the five key ethical dimensions proposed throughout the article: sentience (S), intentionality (I), innocence (N), irreversibility (R), and damage (D).

Each actor within the dilemmas was assigned values ranging from 0 to 1 (or 0 or 1 for binary inputs), and the valence of each action was computed using the ethical balance equation:

$$v = \alpha S + \beta I + \gamma N - \delta R - \epsilon D$$

The dr*ift* rate (μ) for the simulated decision process was calculated as the difference between the valences of Option A and Option B:

μ = VA - VB

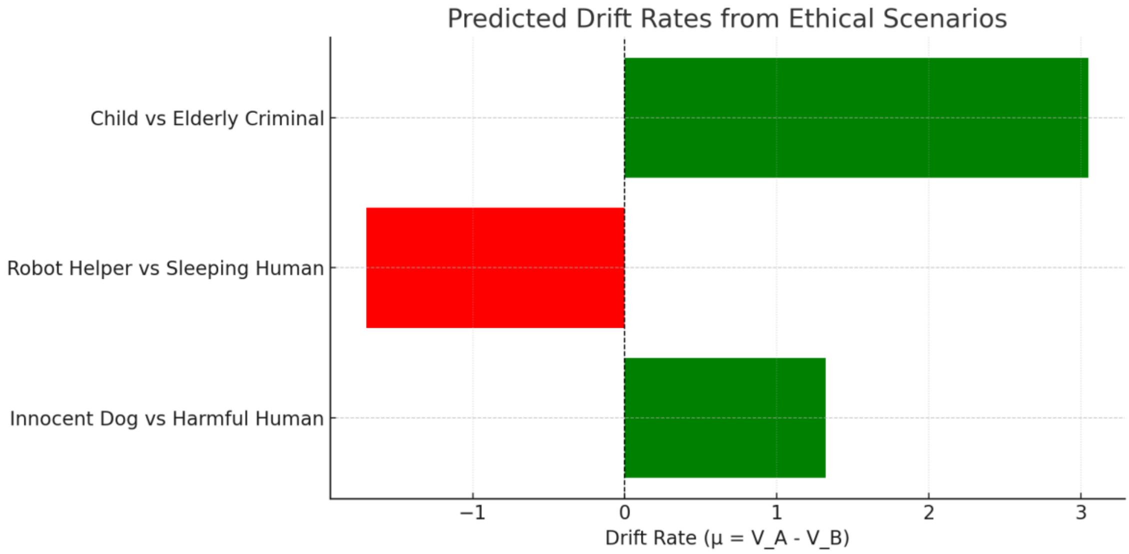
The drift rate serves as an indicator of the expected direction and confidence of the decision: the larger its magnitude, the more likely and rapid the selection of the favored option.

Three dilemmas were tested:

1. Innocent Dog vs. Harmful Human
2. Robot Helper vs. Sleeping Human
3. Child vs. Elderly Criminal

The simulations revealed consistent and interpretable results: scenarios with greater ethical asymmetry produced high drift rates, favoring swift decisions; ambiguous scenarios produced drift rates near zero, indicating internal conflict or indecision. The figure below summarizes the drift rate predictions across the simulated cases.

Figure A1. Predicted drift rates (μ) across three moral dilemmas using the Ethical Balance. Green bars indicate a positive drift rate, meaning Option A is favored by the model's ethical balance. Red bars indicate a negative drift rate, favoring Option B. The length of each bar represents the strength of the decision signal: higher absolute values reflect greater ethical asymmetry and predict faster, more confident decisions, while values near zero indicate moral ambivalence and likely hesitation.



CONCLUSION

This article has proposed a novel framework for addressing moral dilemmas by integrating philosophical ethics with computational modeling. Specifically, it introduces an ethical decision-making algorithm grounded in the Drift Diffusion Model (DDM), allowing moral judgments to be simulated as processes of evidence accumulation shaped by quantifiable ethical dimensions—sentience, intentionality, innocence, damage, and irreversibility. Rather than resolving normative debates, the model aims to illuminate how moral cognition might operate in real-time, particularly under uncertainty or emotional interference. The core contribution lies not in claiming algorithmic supremacy over human ethics, but in formalizing how moral agents might weigh competing values when forced to act. By connecting moral intuition to measurable cognitive dynamics, this framework offers a potential bridge between neuroscience, psychology, and normative philosophy. Moreover, it opens avenues for empirical validation through behavioral experiments and computational simulations. Ultimately, this approach suggests that science can help us better understand—not dictate—our ethical intuitions. It offers a generative model of moral conflict, capable of clarifying hidden assumptions, revealing cognitive biases, and informing the design of ethically sensitive artificial systems. As such, it contributes not only to the modeling of moral cognition, but to the broader interdisciplinary dialogue on how humans—and machines—ought to decide.

Acknowledgments

The author declares no acknowledgments.

REFERENCES

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, *2*(11), 1032-1037. https://doi.org/10.1038/14833

Brink, D. O. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511624612

Bunge, M. (2000). *La investigación científica: Su estrategia y su filosofía*. Siglo XXI.

Bunge, M. (2010). *Matter and Mind: A Philosophical Inquiry*. Springer Science & Business Media.

Bunge, M. (2013). *Pseudociencia e ideología*. Laetoli.

Bunge, M. (2018). *La ciencia: Su método y su filosofía*. Laetoli.

Chan, H.-P., Hadjiiski, L. M., & Samala, R. K. (2020). Computer-Aided Diagnosis in the Era of Deep Learning. *Medical physics*, *47*(5), e218-e227. https://doi.org/10.1002/mp.13764

Chan, H.-P., Samala, R. K., Hadjiiski, L. M., & Zhou, C. (2020). Deep Learning in Medical Image Analysis. *Advances in experimental medicine and biology*, *1213*, 3-21. https://doi.org/10.1007/978-3-030-33128-3_1

Chung, H.-K., Alós-Ferrer, C., & Tobler, P. N. (2021). Conditional valuation for combinations of goods in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1819), 20190669. https://doi.org/10.1098/rstb.2019.0669

Cornish, A., Wilson, B., Raubenheimer, D., & McGreevy, P. (2018). Demographics Regarding Belief in Non-Human Animal Sentience and Emotional Empathy with Animals: A Pilot Study among Attendees of an Animal Welfare Symposium. *Animals : an Open Access Journal from MDPI*, *8*(10), 174. https://doi.org/10.3390/ani8100174

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1346), 1413-1420. https://doi.org/10.1098/rstb.1996.0125

Dawkins, R. (1989). *The Selfish Gene*. Oxford University Press.

De los Ríos-Uriarte, E. (2017). The Question of Method in Clinical Bioethics: Approach to a Methodology Adapted to the Context of Mexican Reality. *Persona y Bioética*, *21*(1), 92-113. https://doi.org/10.5294/pebi.2017.21.1.7

de Waal, F. B. M. (2021). How animals do business. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1819), 20190663. https://doi.org/10.1098/rstb.2019.0663

Earp, B. D. (2015). *La science ne peut pas déterminer les valeurs humaines*.

Feyerabend, P. (1987). *Farewell to Reason*. Verso.

Garrigan, B., Adlam, A. L. R., & Langdon, P. E. (2016). The neural correlates of moral decision-making: A systematic review and meta-analysis of moral evaluations and response decision judgements. *Brain and Cognition*, *108*, 88-97. https://doi.org/10.1016/j.bandc.2016.07.007

Goldman, A. D., & Becerra, A. (2024). A New View of the Last Universal Common Ancestor. *Journal of Molecular Evolution*, *92*(5), 659-661. https://doi.org/10.1007/s00239-024-10193-w

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, *44*(2), 389-400. https://doi.org/10.1016/j.neuron.2004.09.027

Harris, S. (2011). *The Moral Landscape: How Science Can Determine Human Values*. Simon and Schuster.

Heyes, C. (2018). Empathy is not in our genes. *Neuroscience and Biobehavioral Reviews*, *95*, 499-507. https://doi.org/10.1016/j.neubiorev.2018.11.001

Jonsen, A. R. (with Siegler, M., & Winslade, W. J.). (1982). *Clinical ethics, a practical approach to ethical decisions in clinical medicine*. Macmillan.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A Flaw in Human Judgment*. Little, Brown.

Knobe, J. (2005). Theory of mind and moral cognition: Exploring the connections. *Trends in Cognitive Sciences*, *9*(8), 357-359. https://doi.org/10.1016/j.tics.2005.06.011

Krämer, W. (2014). Kahneman, D. (2011): Thinking, Fast and Slow. *Statistical Papers*, *55*(3), 915-915. https://doi.org/10.1007/s00362-013-0533-y

Kuhn, T. S. (1994). *The structure of scientific revolutions* (2. ed., enlarged, 21. print). Univ. of Chicago Press.

Marewski, J. N., & Gigerenzer, G. (2012). Heuristic decision making in medicine. *Dialogues in Clinical Neuroscience*, *14*(1), 77-89. https://doi.org/10.31887/DCNS.2012.14.1/jmarewski

Mendez, M. F. (2009). The Neurobiology of Moral Behavior: Review and Neuropsychiatric Implications. *CNS spectrums*, *14*(11), 608-620. https://doi.org/10.1017/s1092852900023853

Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An Investigation of Moral Judgement in Frontotemporal Dementia. *Cognitive and Behavioral Neurology*, *18*(4), 193. https://doi.org/10.1097/01.wnn.0000191292.17964.bb

Miller, M. B., Sinnott-Armstrong, W., Young, L., King, D., Paggi, A., Fabri, M., Polonara, G., & Gazzaniga, M. S. (2010). Abnormal Moral Reasoning in Complete and Partial Callosotomy Patients. *Neuropsychologia*, *48*(7), 2215-2220. https://doi.org/10.1016/j.neuropsychologia.2010.02.021

Moll, J., & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive Sciences*, *11*(8), 319-321. https://doi.org/10.1016/j.tics.2007.06.001

Moretto, G., Làdavas, E., Mattioli, F., & di Pellegrino, G. (2010). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *Journal of Cognitive Neuroscience*, *22*(8), 1888-1899. https://doi.org/10.1162/jocn.2009.21367

Myers, C. E., Interian, A., & Moustafa, A. A. (2022). A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences. *Frontiers in Psychology*, *13*, 1039172. https://doi.org/10.3389/fpsyg.2022.1039172

Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: The utility of a cognitive approach. *Neuroscience and Biobehavioral Reviews*, *29*(3), 469-491. https://doi.org/10.1016/j.neubiorev.2005.01.002

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews. Neuroscience*, *9*(2), 148-158. https://doi.org/10.1038/nrn2317

Peters, J., & D'Esposito, M. (2020). The drift diffusion model as the choice rule in inter-temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and controls. *PLoS Computational Biology*, *16*(4), e1007615. https://doi.org/10.1371/journal.pcbi.1007615

Richens, J. G., Lee, C. M., & Johri, S. (2020). Improving the accuracy of medical diagnosis with causal machine learning. *Nature Communications*, *11*, 3923. https://doi.org/10.1038/s41467-020-17419-7

Roberts, I. D., & Hutcherson, C. A. (2019). Affect and Decision Making: Insights and Predictions from Computational Models. *Trends in Cognitive Sciences*, *23*(7), 602-614. https://doi.org/10.1016/j.tics.2019.04.005

Santens, P., Vanschoenbeek, G., Miatton, M., & De Letter, M. (2018). The moral brain and moral behaviour in patients with Parkinson's disease: A review of the literature. *Acta Neurologica Belgica*, *118*(3), 387-393. https://doi.org/10.1007/s13760-018-0986-9

Seymour, B., & Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron*, *58*(5), 662-671. https://doi.org/10.1016/j.neuron.2008.05.020

Singer, P. (1972). Famine, Affluence, and Morality. *Philosophy and Public Affairs*, *1*(3), 229-243.

Thompson, B., & Griffiths, T. L. (2021). Human biases limit cumulative innovation. *Proceedings. Biological Sciences*, *288*(1946), 20202752. https://doi.org/10.1098/rspb.2020.2752

Whelehan, D. F., Conlon, K. C., & Ridgway, P. F. (2020). Medicine and heuristics: Cognitive biases and medical decision-making. *Irish Journal of Medical Science*, *189*(4), 1477-1484. https://doi.org/10.1007/s11845-020-02235-1

Wu, Y. E., & Hong, W. (2022). Neural basis of prosocial behavior. *Trends in Neurosciences*, *45*(10), 749-762. https://doi.org/10.1016/j.tins.2022.06.008