„Zoophilologica. Polish Journal of Animal Studies"
Nr 2 (14)/2024, pp. 1–43
ISSN 2451-3849
DOI: https://doi.org/10.31261/ZOOPHILOLOGICA.2024.14.12

Zbigniew Słuszkiewicz
https://orcid.org/0000-0001-7734-8484
Uniwersytet im. KEN w Krakowie

Of Rats and Men I: A Pragmatist Take on the Concept of Free Will as a Challenge to the Human-Animal Dichotomy

О крысах и людях I – Подход прагматика к концепции свободы воли как вызов для дихотомии человек–животное

Абстракт

Статья является первой частью триптиха о субъектности. В ней проводится критический анализ тех аргументов, основанных на понятии свободной воли (СВ), которые наиболее часто вновь возникают в обсуждениях внеличных умов в контексте дихотомии «человек-животное». Включив исследование в рамки прагматической философии, поддержанной широкими познаниями в когнитивных науках, автор утверждает, что эта сеть утверждений неубедительна, поскольку онтология СВ: 1) не согласуется с текущими эмпирическими данными; 2) представляется поверхностной конструкцией, не отражающей сложность процессов принятия решений человеком; и 3) не имеет значения для опыта. СВ, вместо того чтобы быть уникальной способностью ума, раскрывается как антропоцентрический и WEIRD-созданный артефакт, не обладающий объяснительной

Of Rats and Men I: A Pragmatist Take on the Concept of Free Will as a Challenge to the Human-Animal Dichotomy

Abstract

The article is the first part of a triptych on agency. It critically analyses those arguments drawing from the notion of free will (FW) that most frequently resurface in discussions of non-human minds within the context of the human-animal divide. By embedding the investigation within pragmatic philosophy, backed by broadly understood cognitive sciences, the author argues that this web of assertions is unconvincing, as FW's ontology: 1) is inconsistent with current empirical knowledge; 2) appears to be a superficial construct, failing to reflect the intricacies of human decision-making processes; and 3) is inconsequential for experience. Rather than being a unique capability of the mind, FW reveals itself as an anthropocentric and WEIRD-made artefact, which lacks explanatory power regarding human behaviour. As such, when applied to non-human species, it qualifies as anthropofabulation. Consequently, the author proposes replacing FW

силой в отношении человеческого поведения. Таким образом, при применении к внеличным видам оно квалифицируется как антропофабуляция. Следовательно, автор предлагает заменить СВ эмпирически обоснованной концепцией субъектности, выведенной из принципа свободной энергии Карла Фристона и поддержанной выводами процессуальной философии биологии. Такая модель позволяет выявить природу субъектности в её продвинутой форме, которая, возможно, разделяется по крайней мере некоторыми другими видами. Это, в свою очередь, способствует созданию градуируемых концепций моральной субъектности, подобных категории «морального субъекта» Марка Роуландса. В последующих статьях данная тема будет обсуждаться более подробно.

Ключевые слова: Прагматизм, Свобода воли, Дихотомия человек-животное, Антропофабуляция, Марк Роуландс

with an empirically informed concept of agency derived from Karl Friston's free-energy principle, supported by insights from the processual philosophy of biology. Such a framework allows for capturing the nature of agency in its advanced form, plausibly shared with at least some other species. It thus facilitates the construction of gradable concepts of moral agency, in the likes of Mark Rowlands's category of "the moral subject." Subsequent articles will discuss this subject matter further.

Keywords: pragmatism, Free Will, human-animal dichotomy, anthropofabulation, Mark Rowlands

> [From the] Cartesian point of view, the genius of pragmatism is to get all of its explanatory priorities backwards, [it] is Cartesianism read from right to left. Pragmatism is in the very air that our cognitive science breathes (…) the bad cold that we've all come down with. It's what must be overcome, preferably by next Tuesday.[1]
>
> Jerry Fodor

> [E]vidently, humans prefer to consider themselves the lowest of angels rather than the highest of beasts. I am of the opinion that if we were to comprehend fully the awesome nature of mindness, we would, in fact, respect and admire each other all the more.[2]
>
> Rodolfo Llinás

Introduction

In the uncharted waters of philosophical discourse, explorers of the metaphysical depths had navigated the swirling currents of controversies flooding the concept of

[1] Jerry Fodor, *Lot 2: The Language of Thought Revisited* (Oxford, New York: Clarendon Press, 2008), 11–12.

[2] Rodolfo Llinas, Introduction to *I of the Vortex: From Neurons to Self*, Reprint edition (Cambridge, Mass. London: Bradford Books, 2002).

free will (FW) for almost two millennia. The ontological status of FW, its properties, its necessary and sufficient conditions, the impact it has on society and morals, the uniqueness it imparts to our species among living beings, and, of late, its compatibility with a scientifically informed, deterministic worldview – these motifs were meant to serve as lighthouses, illuminating and guiding discussions. It has been so since the concept's formation, even though, in its metaphysical guise, the idea of FW resembles an elusive shape of kilwater foam on the sea surface, endlessly trailing the strayed flagship of perennial philosophy.

In the last decades of the 20th century, however, the course of the debate has shifted dramatically. The seminal work of Benjamin Libet, a *protégé* of the famous dualistic neuroscientist John Eccles, significantly propelled this change. Libet explored the temporal relationship between the conscious intention to act and the timing of brain activation (readiness potential). While attempting to prove the downward causal powers of mental states over brain activity, the future Nobel prize laureate discovered, much to his chagrin, that causal relations between intention to move and neuronal activity appear to flow in the opposite "bottom-up" direction.[3] From that point, stormy wranglings over Libet's results and the subsequent experimental paradigms have reinvigorated the scientific and philosophical communities. However, we shall not dive into the details of this conceptual and methodological abyss here.[4] It suffices to say that the traditionally accepted ontological status of the FW, seemingly anchored in both theological considerations and common-sense impressions,[5] has been challenged by empirical findings, sowing enough doubts to resume inquiry into the meaning of this concept. Despite the experiential investiga-

[3] Benjamin Libet et al., "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness – Potential). The Unconscious Initiation of the Freely Voluntary Act," *Brain*, 106, no. 3 (1983): 623–42, accessed June 3, 2023, https://doi.org/10.1093/brain/106.3.623.

[4] For an up-to-date overview of "the Libet debate," see Robert Sapolsky, *Determined: A Science of Life Without Free Will* (London: The Bodley Head, 2023), 18–41. The article will not engage in it. We agree with Sapolsky, who himself finds the opposite parties' emphasis on the importance of Libet's paradigm unnecessary and myopic, comparing it to the attempt to reconstruct a movie after watching its last scene. Most importantly, focusing on the causal relationship between the urge to move and the awareness of the decision does not explain the intentions' origins. Suppose one tries to prove the existence of FW from a naturalistic vantage without referring to dualistic intuitions. In that case, one faces the impossible task of showing that before the decision-forming process, without any outside influence on its shape or structure (electrical, chemical or other) at the moment of intent formation, at least a single neuron has activated itself – a task which Sapolsky called "uncaused cause challenge."

[5] For the origins of the FW concept, see Michael Frede, *A Free Will: Origins of the Notion in Ancient Thought*, 1st ed. (Berkeley: University of California Press, 2011). For the analysis of the conceptual relations between scientific and common sense (folk psychology) approaches, see Pim Haselager, "Conceptual Revisions. Intentions and Free Will in the Light of Cognitive Neuroscience," in *Scientific Challenges to Common Sense Philosophy*, eds. Rik Peels, Jeroen de Ridder, and Rene van Woudenberg (New York/London: Routledge: Taylor & Francis Group, 2020), 104–20.

tions into the guts of the compass that was once thought to indicate FW's bearings overflowed the philosophical debate and rocked its conceptual balance, leading the helmsmen of the traditional course toa kind of intellectual seasickness, they continue to cling to the wheel.[6] The prototypical defense of the FW and its indispensability has been encapsulated in an elegant (albeit slightly confusing) manner by the esteemed cognitive philosopher and cleric Józef Bremer in his introduction to an extensive overview of the current state of the controversy. We will refer to it as "The Free Will Conundrum" (FWC):[7]

> Free will is one of the person's attributes. Due to its possession, we are able to perform independent voluntary decisions and make choices ranging from immediate and straightforward to long-term or complex moral ones. The free will talk is not about something marginal but the essence of our and others' self-knowledge. This capability is the bedrock for our flourishing in society, the foundation of our social, judiciary, political, and moral systems, where it is inherently presumed to varying degrees. Since free will is entwined with personal accountability for one's actions, it would be next to impossible to talk about good and evil deeds, punishment and reward, or praise and blame without it. The ability to make free decisions *distinguishes us from animals*, whose conduct is governed by instincts and reflexes. We take for granted that without free will, we humans would be mere *automatons*. Every single action performed by us, every moral transgression, would be justified by our inability to do otherwise, given that our brains were determined by the laws of chemistry and physics from the outset, even before our consciousness had the opportunity to confront the brain's doings critically and ultimately to resist them. Nonetheless, to what extent are we genuinely free in our choices and decision-making? How constrained are we by biophysical realities, worldly structures, and social rules? The problem of free will, a prominent topic throughout history, still prevalent in modern contemplations, arises at the intersection of two contrasting convictions: our subjective, personal folk experience of freedom and our adherence to objective, scientific beliefs about a world governed by deterministic forces.[8]

[6] We apologise to the Reader for this pretentious maritime metaphor, hoping our reasons for it will become progressively transparent.

[7] The choice of such a name results directly from the antinomy hidden in the quote. If we admit that mental life results directly from brain activity, then on what basis do we claim that consciousness can resist it? We emphasise that placing free will in Kant's intelligible world does not solve the problem; it only shoves it into the transcendental domain.

[8] Translations and italics are mine. The quote above is from: Józef Bremer, *Czy wolna wola jest wolna?* (Kraków: Wydawnictwo WAM, 2013), 9.

For brevity, we must forgo the analysis of the author's juxtaposition of experienced personal freedom and determinism as a (Sellarsian in spirit) opposition between commonsense and scientific understandings of the world.[9] Instead, we are looking to focus on one of his premises.

Bremer aptly points out FW's potential limitations discussed in the classical discourse. Nevertheless, he takes it as given – as a defining attribute of human beings, the cradle of social order, and the source of genuine moral responsibility. Concomitantly, conforming to longstanding philosophical tradition, he bolsters a boundary between humans and other animals by using this *ex hypothesi* capacity to clear-cut a demarcation line. Some might say that a quoted conviction represents one person's view on the matter. Regrettably, this is not the case. The very fact that the Cartesian premise is included in the introduction to an overview publication suggests that it is considered uncontroversial, and as such, it will not be discussed in the following chapters. In other words, the belief that the lack of free will would make us *brute automatons* continues to be a mainstream view in philosophy and, in this form, penetrates common background knowledge.

However, to say that the author of FWC oversimplifies the cognitive abilities of animals by narrowing down the explanation of their behaviour to the confinements of instincts and reflexes would be a gross understatement in the face of the advancements in the natural sciences over the past decennaries. The disregard for this "nuance," on its face, would call for a more pragmatic exploration of whether FW is, in fact, a defining trait of our species, what its societal implications are, and what its potential relationship with the decision-making processes of cognisers might be, so that we could decide whether this category is a proper departure point for separating humans from other animals.[10] But before we proceed, we have to take a step back to sketch the context that motivates us to investigate the pointed issue and perspective we embrace.

[9] See Wilfrid Sellars, "Philosophy and the Scientific Image of Man," in *Science, Perception and Reality* (Atascadero: Ridgeview Publishing Company, 1991).

[10] One might counter our take on this issue on the grounds of the "person" concept. However, aside from the equally metaphysically rooted (albeit useful) category of "dignity," no conceptual obstacles prevent the attribution of minimal personhood to some animals (Barbara Tomczyk, *Podmiotowość rozszerzonych systemów poznawczych* (Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, 2022, 27). Unfortunately, this paper does not have the scope to delve into a discussion on metaphysical personhood concerning animals (for discussion, see Mark Rowlands, *Can Animals Be Persons?* (New York: Oxford University Press, 2019).

Zbigniew Słuszkiewicz

The Pragmatist Issue with Metaphysical Concepts

Following its emergence in the late nineteenth century, classical American pragmatism, in all its heterogeneity, consistently shares three main philosophical themes: it underscores the active and subjective role of interpretative practices in knowledge acquisition (inquiry), be they scientific or otherwise, it holds the fallibilist stance towards the nature of knowledge, by indicating the fragility of human experience; and it stoutly rejects any foundationalist accounts of truths as the fruitless pursuits of certainty. The above philosophical commitments follow, in part, from the pragmatic doubts about the central tenet of the perennial philosophy: its fixation on the commonsensical-sounding view about the reality of physical objects as existing independently of the mind. This conviction, in turn, has led classical philosophers to embrace the correspondence theory of truth – a claim that beliefs of the mind can accurately represent (mirror) the world – and to the epistemological thesis that the mind's knowledge rests on firm conceptual foundations.

Reservations for these metaphysical flagposts situate pragmatism in the anti-Cartesian philosophical camp, while keeping it within a tenable distance from pure rationalist and empiricist views.[11] From the philosophical standpoint, pragmatic moderate skepticism has led to the emergence of a new approach to epistemology by transforming its traditional obsession with certainty to a kind of intellectual, empirically informed exercise committed not to merely abstract speculation but to resolving real-life problems – "the aim is to understand how things, in the broadest possible sense of the term, hang together, in the broadest possible sense of the term."[12] This remark, made by Wilfrid Sellars, brings us back to the subject of the article – the concept of free will and its role in our understanding of ourselves and our relations to the rest of the living.

Since its inception, this category has rested upon dualistic intuitions about the mind (rational soul) and the world, serving as a metaphysical opposition to the notion of determination. In a similar vein, the concepts of humanness and non-humanness were supposed to represent some true ontological distinction. Although, from the pragmatist's vantage point, concepts are human-made epistemological devices which are "true" as long as they fulfil their functions. There are no "metaphysical Truths" to them – or, to put it in Jamesian terms, concepts are true insofar as they entail experiential consequences for those who take them to be

[11] So far, advances in modern cognitive and natural sciences regarding human cognition and physical reality have mostly testified in favour of the pragmatist stance.

[12] Wilfrid Sellars, "Philosophy and the Scientific Image of Man," in *Frontiers of Science and Philosophy*, ed. Robert Colodny (Pittsburgh: University of Pittsburgh Press, 1962), 37.

real.[13] Thus, guided by the pragmatic approach, we could immediately ask three questions: (1) Does the concept of FW help us comprehend the human condition, and if so, how does it work? (2) Does it help in understanding other animals, and (3) What role does it play in our image of the world and us in it?

Starting from the last one, if FW were inconsequential on every level, we could leave it as it is and not bother. However, as we have seen in the case of FWC, it would not be an accurate description of the situation by a long shot. The issue at hand is more than merely academic trivia, as it has tangible consequences for other species. Some societies' policy-makers overtly invoke these kinds of views to justify our exploitative policies towards them. Such attitudes are expressed in derogatory references to nonhuman animals (from here on referred to as animals) as "bleating meat sacks" or in framing animal rights movements as "attempts to animalise man and humanise animals."[14] Indeed, for many thinkers throughout history, the notion of FW marked the sharp distinction between "the humane" and "the inhumane."[15]

In the pragmatist, more liberal depiction, FW is the crucial device with which philosophers draw the demarcation line between animals and us, but detrimental consequences follow this procedure. The problem becomes apparent when we look at the recently emerging field of research on animal morality.[16] Here lies our main inspiration for picking up those questions and expanding our deliberations on three related papers. It started with the idea introduced by cognitive philosopher and embodied cognition theorist Mark Rowlands, who challenged the classically understood, anthropocentric view on the phenomenon of morality as a distinctively human phenomenon.

In *Can Animals Be Moral?* Rowlands argues that there are no convincing philosophical arguments against introducing an intermediate category of moral subjectivity and placing it between the acknowledged categories of the full moral agent and moral patient.[17] A being classified as "the moral subject" would count as such based on her ability to carry out actions with moral characteristics. Her deeds would be caused by the emergence of affective attitudes that fulfil motivational functions (as the building blocks of basic moral emotions). A subject of this type would differ from a human agent (a subject capable of maintaining a consistent moral attitude)

[13] See William James, *Odmiany Doświadczenia Religijnego. Studium Ludzkiej Natury*, trans. Jan Hempel (Warszawa: Aletheia, 1902/2011), 438–464.

[14] See *Tadeusz Guz – Wykład*, 2018, accessed March 10, 2023, https://www.youtube.com/watch?v=1_LqnK2j_Do.

[15] See Emile Benoit, *Beasts in Eden: The Humane and the Inhumane* (San Diego: Eudaimonia Press, 2016).

[16] We should mention Frans de Waal, who recently passed away, but also Sarah Brosnan, Mark Bekoff, Jessica Pierce, Kristin Andrews, Susana Monsó, and Judith Benz-Schwarzburg. We apologise for omitting many other philosophers and scientists working in this rapidly evolving field.

[17] Mark Rowlands, *Can Animals Be Moral?*, reprint edition (Oxford University Press, 2012).

in the path by which she reaches the postulated status. It would follow from the creature's ability to track morally relevant aspects of the scene accompanied by the execution of acquired moral skills shaped by mastering the particularities of her species' social rules. In this approach, specific emotions, dependent on context and experience, evoke mental states with moral content, thereby constituting moral reasons for actions. The occurrence of a given affect pertinent to the context could then be captured as tantamount to a moral assessment leading to a proper behavioural response to the demands of the situation, resulting in a proper skillful goal-oriented action. Such being would gain normative sensitivity by developing social practices through observational learning and mastering habits of conduct in specific, repeated social contexts. Crucially, a morally motivated entity would not have to be able to grasp the (propositional) semantic meaning of moral concepts, nor would she have to possess reflective self-awareness or bear moral responsibility for her actions.

However, here we find a drawback to this proposition. One can always insist that for any creature to act as a genuine moral being, *she has to* be able to bear moral responsibility. As we have seen, for most philosophers, the only route to this category leads *via* the possession of free will. Rowlands rejects this argument, designing a thought experiment on the moral judgment of Hitler's motivations in a deterministic world (these would still be valued as evil).[18] Although compelling, Rowland's conclusion could be insufficient for abandoning the mainstream idea. We could also argue that the whole view of morality as requiring responsibility, judgment, and behaviour working together rests on the *amalgamation fallacy*, as Susana Monsó pointed out.[19] Yet, as apt as it is, in our view, her argument does not fix the main problem behind responsibility demand – the deeply grounded philosophical conviction regarding the human-animal dichotomy. And this is what we want to tackle by looking at the function of the concept of FW as it pertains to humans and animals.

Before moving further, we need to make explicit one more methodological constraint we took. We will not engage in specific philosophical brawls between backers of the various properties of FW or dive in-depth into the metaphysical claims upon which they are built. This paper is not designed to report the state of debate on this topic. Neither do we follow any particular theory of FW. Even though we could have found a worthy defense for the animal version of it,[20] we see such attempts as misguided, contending that the roots of the problem lie elsewhere – in the concept

[18] Mark Rowlands, "Animals as Moral Subjects," in *The Routledge Handbook of Philosophy of Animal Minds*, ed. Kristin Andrews and Jacob Beck (London/New York: Routledge, 2018), 471–72.

[19] Susana Monsó, "Morality and Mindreading in Nonhuman Animals" (Universidad Nacional de Educación a Distancia (UNED), 2016), 54–64, accessed June 23, 2023, http://e-spacio.uned.es/fez/eserv/tesisuned:Filosofia-Smonso/MONSO_GIL_Susana_Tesis.pdf.

[20] See Helen Steward, *A Metaphysics for Freedom* (Oxford: Oxford University Press, 2014).

of FW itself. Instead, we will look at the most common themes recurring in philosophical literature in the animal context. By that, we mean confronting the notions of rationality, "the space of reason," "the second-order desire," "the ability to choose otherwise," and the worry about humanity without the concept of free will.

Due to the breadth of the topic, the article is divided into three separate papers. In all of them, we embrace the framework of modern cognitive pragmatism, as it allows us to supersede traditional conceptual restrictions and dichotomies accompanying considerations about other minds.[21] Most of the sources we will use throughout the papers (but mainly the second and third ones) stem from researchers and philosophers committed to the recently initiated movement in cognitive science dubbed "the pragmatic turn."

Following the pragmatism founder – Charles Sanders Peirce, who criticised the classical (Cartesian) insistence that an argument should resemble a chain only as strong as its weakest link – we embrace Peirce's line of reasoning encapsulated in the metaphor of a cable – while its individual interwoven fibres may be frail, when sufficiently numerous, they merge to form a robust line. This underscores our reliance on multiform argumentation rather than a singular, definitive argument, as well as pragmatic resignation from the pursuit of establishing "infallible truth."[22]

The structure of the first paper is as follows: The opening section is dedicated to investigating explanatory practices of animal behaviour, aiming to illuminate their

[21] Since the dawn of the pragmatist philosophy, it has espoused the evolutionary framework, and its proponents generally defend the Darwinian thesis on cognitive continuity between species "in degree rather than in kind." Therefore, it should come as no surprise that classical pragmatism has become one of the primary conceptual resources (next to phenomenology) for the embodied cognition paradigm and contemporary neuroscience. It is so mainly due to pragmatic (Greek πρᾶγμα, πρᾶγμᾰτος – act, deed, affair, concrete reality) emphasis on the critical input of environmental context for the cognitive process, which manifests itself in goal-oriented actions of the organism regardless of the species it belongs to, see Mark Johnson and Tim Rohrer, "We Are Live Creatures: Embodiment, American Pragmatism and the Cognitive Organism," in *Body, Language, and Mind*, Vol. 1, eds. Tom Ziemke, Jordan Zlatev, and Roslyn Frank (Berlin: Mouton de Gruyter, 2007), 17–54. What follows from the pragmatist heresy are substantial consequences for the philosophy of animal minds; there is, for example, a strand of contemporary pragmatism whose representatives, following Alexander Bain, stand in sharp opposition to the dichotomic over-intellectualised philosophical tradition regarding the propositional nature of beliefs, desires and intentionality, fully recognising that animals possess these mental states, see Aaron Zimmerman, *Belief: A Pragmatic Picture* (Oxford: Oxford University Press, 2018).

[22] The fallibilist doctrine lies at the core of Peirce's philosophical project. In contrast to Cartesian tradition, he viewed all human knowledge as fragile and uncertain. Uncertainty and efforts dedicated to resolving it are, in turn, the necessary antecedents of any knowledge; see Alexander Klein, "Review of Peirce's Pragmatic Theory of Inquiry: Fallibilism and Indeterminacy, by Elizabeth Cooke," *Notre Dame Philosophical Reviews*, 11 October 2007, accessed June 4, 2023, https://ndpr.nd.edu/reviews/peirce-s-pragmatic-theory-of-inquiry-fallibilism-and-indeterminacy/. The subsequent paper will expand on this intuition in the modern context.

shortages and exposing tacit assumptions influencing tendencies for using those as justifications for human exceptionalism. After that, we tackle recurring themes in animal mind philosophy invoking properties of the putative FW used as a mark of ontological difference. This includes touching upon the related category of rationality and reflecting on the classic hallmark criteria of FW. In the next section, we employ the pragmatic maxim to discern the influence that belief in FW may have on human conduct. We intend to demonstrate that the FW category, as traditionally understood, does not characterise humans and, as such, cannot serve as a justification for the separation of humans from other animals because, from the pragmatic vantage, when applied to other species, it meets the definition of anthropofabulation. We end the article with a short philosophical conclusion, pointing out the possibility of looking at the phenomenon termed "free will" from a broader perspective and without the necessity of its usage. The suggested proposition (including a bottom-up approach) does not exclude other animals. At the same time, it indicates that the sense of agency (SoA), the source of FW intuitions, can be shared to a degree with sophisticated enough acting creatures.

In the ensuing paper "Of Rats and Men II: A Pragmatist Reconstruction of the Basis of Agency *via* the Free-Energy Principle," we introduce the pragmatically inspired free-energy principle (FEP) and advocate for the interpretation through its lenses of the properties, which are thought of as constitutive for the free will. We aim to reconstruct a scope of the biological foundations of agency through one of the FEP's corollary theories (predictive processing, PP) as it applies both to humans and other animals. Additionally, we will substantiate our reconstruction with insights derived from the modern processual approach in the philosophy of biology. We employ the pragmatic maxim once again, this time on a higher theoretical level. That means we will not engage in discussion about the realism of the FEP. Instead, we intend to investigate how the agency originates without the FW notion, assuming the FEP's explanatory framework to be on the right track.

In the final part of the triptych, "Of Rats and Men III: A Pragmatist Reconstruction of Advanced Agency *via* the Active Inference," we will examine the characteristics of sophisticated forms of agency from the perspective of the second theory entailing FEP. Namely, we intend to explore advanced agency through the lens of active inference (AIN). Next, the most consequential features accompanying its emergence (counterfactual inference and psychophysiological self-evidencing) will be juxtaposed with the empirical results obtained with rodents to justify the claim that these animals may be capable of experiencing agentive activity at this level. To further investigate the subjective aspect of agency in animals, we put forth a usage of the empirical experimental ("the temporal binding" paradigm) design, suggesting its interpretation and philosophical formulation can be grounded on the active inference's (AIN) perspective on the agency. In our three-part-article for illustrative

purposes, we have chosen rodents as exemplars for our argument. Although we could have picked some lower-hanging fruits (e.g. primates),[23] our choice was driven by the intent to curtail the tendency for direct anthropomorphic (pedomorphic) comparisons and to extend the moral discussion beyond our closest evolutionary relatives. The more general thesis we argue for is that similarities in animal and human capacities (including apparent FW properties) could be interpreted as originating from the same life function shaped by a particularities of ecological *niche* inherited by the given species. In the triptych, we contend that a) explanations of cognitive capacities appealing to evolutionary parsimony (closeness) are too narrow, and b) that if a notion of well-developed (sophisticated) agency can be reconstructed without retracting to philosophical over-intellectualised constructs, then it is possible to establish gradual categories of moral agency, such as Rowlands'.

When "Free-Floating Rationales" Meets the "Mindless Drivers"

Bremers' introduction to his overview of the free will debate, which we summarised as FWC, accurately captures classical intuition that, while we perceive ourselves as residing in the Aristotelian "space of reasons," we grant animals merely the necessity of following instincts. This is where the gap between "us and them" resides. Is it justified, however? We begin addressing this conjecture from where it starts.

The presumption that being bereft of FW would make automatons out of us seems to allude to the Cartesian tradition in accepting as a self-evident and unquestionable fact that the entire functioning of animals in the world comes down to the automatic triggering of drive mechanisms. Those drives are what classical philosophers and lay people commonly refer to as "instincts." Thus, in the scientific disguise, the ontological difference which Bremer suggests manifests itself in the worry that depriving man of FW would relegate him to the animal dimension of the

[23] Even capuchin monkeys seem to experience a sense of agency (SoA) and perhaps even believe in the limitlessness of the will (see Sapolsky, *Determined*, 242). For example, they prefer the choices they have already made, even if they were forced to make them, and they can also distinguish instances in which humans wanted to help them but could not from situations in which they could help but did not want to, see Louisa Egan, Paul Bloom, and Laurie Santos, "Choice-Induced Preferences in the Absence of Choice: Evidence from a Blind Two Choice Paradigm with Young Children and Capuchin Monkeys," *Journal of Experimental Social Psychology*, 46, no. 1 (2010): 204–7, accessed June 15, 2023, https://doi.org/10.1016/j.jesp.2009.08.014; Webb Phillips et al., "'Unwilling' versus 'Unable': Capuchin Monkeys' (*Cebus Apella*) Understanding of Human Intentional Action," *Developmental Science*, 12, no. 6 (2009): 938–45, accessed June 15, 2023, https://doi.org/10.1111/j.1467-7687.2009.00840.x.

involuntary commencing of genetically encoded programs for behaviour prompted by environmental stimuli. At this point, we need to briefly tackle the key notion on which the above concern rests.

First of all, it should be noted that the concept of "instinct," both in folk psychology and by philosophers, is frequently overused, despite the fact that its explanatory power for empirical research has always been illusory; it is an ill-defined and vague term.[24] Second, "instincts" fail to encompass a wide array of animal cognitive abilities, such as – to scratch the surface – flexible, goal-oriented behaviours, self-control, empathy, social learning, generalisations, insight, tool-making and tool-use, rule-based categorisation, counterfactual reasoning or inferential processes.[25]

Given the pragmatist approach to meaningless concepts, it should not be a stunner that invocations to instinct as an explanation (one to rule them all) were already emphatically criticised by John Dewey over a century ago as an oversimplification of the myriad factors contributing to action.[26] Pragmatic critique has proven to be on track, as the debate over the explanatory utility of "instinct" is, for the most part, settled nowadays. For example, Robert Sapolsky did not devote even a paragraph to elucidate this concept in his thorough analysis of various evolutionary, biological, and societal facets that converge to form a single decision-making moment.[27] Rather, it is now argued that continuity exists between genetically inherited fixed response patterns (instincts) and flexible goal-oriented behaviour

[24] Mark Blumberg, "Development Evolving: The Origins and Meanings of Instinct," *Wiley Interdisciplinary Reviews. Cognitive Science*, 8, no. 1–2 (2017), accessed June 3, 2023, https://doi.org/10.1002/wcs.1371.

[25] See Frans de Waal, *Bystre zwierzę. Czy jesteśmy dość mądrzy aby zrozumieć mądrość zwierząt?*, trans. Łukasz Lamża (Kraków: Copernicus Center Press, 2019); Michael Beran, *Self-Control in Animals and People*, 1st edition (London: Academic Press, 2018); Ksenia Meyza and Ewelina Knapska, *Neuronal Correlates of Empathy: From Rodent to Human* (Cambridge, Mass: Academic Press, 2018); Christian Keysers, *Empatia. Jak odkrycie neuronów lustrzanych zmienia nasze rozumienie ludzkiej natury*, trans. Łukasz Kwiatek (Kraków: Copernicus Center Press, 2020); Judith Benz-Schwarzburg, *Cognitive Kin, Moral Strangers? Linking Animal Cognition, Animal Ethics & Animal Welfare* (Leiden: Brill, 2020).

[26] John Dewey, *Human Nature and Conduct* (Overland Park: Digreads.com Publishing, 1922/2012: 60–63). The history of the critique of the meaning of this notion can be traced back, at least, to David Hume, who observed that "instinct" often serves as a dummy placeholder for explanations. This makes it harder to understand our reasoning capacities, which we assume are of different origins. For Hume, though, both modes of functioning in nature originate in the same natural phenomenon – *habit*, so one might just as validly assert that human beings possess the "instinct of reason." It is ironic that, for the most famous sceptic of the Enlightenment, nothing was more evident than that animals and men exhibit analogous capacities for reasoning, see David Hume, *Traktat o naturze ludzkiej*, trans. Czesław Znamierowski (Warszawa: ALETHEIA, 2023), 250–53.

[27] See Robert Sapolsky, *Behave: The Biology of Humans at Our Best and Worst* (London: Vintage, 2018).

(rationality).[28] This does not lead to the necessity of negating the existence of some built-in, predetermined, genetically ingrained developmental pathways underlying "extended phenotypes" which activate in the face of adequate environmental cues.[29] However, in no way are we able to paint animal behaviour in its entirety with such brushes. Moreover, even some of the mechanisms termed "instincts" often turn out to be a form of context-dependent learning, and their heredity is a fact about *Homo sapiens* as well.[30] The distinction when it comes to human "instincts" or "intuitions" seems, for the most part, to rest on our socio-cultural practices crafted to justify these; we will expand upon this claim later in the paper.[31]

A charitable suggestion for adherents of the *automatons* presupposition adduced to in the FWC might be to replace the outdated term in question with concepts derived from behavioural strategies (stimulus-response or associative models) in which instinctual reflexes can be captured as "unconditional responses."[32] *Prima facie*, this proposition could serve as a persuasive alternative, given that within those conceptual frames, in line with the philosophical *praxis*, animals are perceived as passive objects that only respond (stiffly react) to incoming environmental determinants. This non-cognitive "learning" process can be successfully mimicked in the lab by repeatedly nudging animals with stimuli – an artificial procedure (deprived of ecological context) aiming at shaping their behaviour through conditioning.

However, the ability of organisms to reorganise their environment and themselves to match their goals becomes bewildering if they are to be considered mere stimulus-response machines. While the production of approximated behavioural sequences may be regarded as an experimental success, more often than not, animals display a creative bypassing of the researchers' conjectures. This flexibility seems to be the norm pithily expressed in the aphorism known as the Harvard Law

[28] Roberto Maffei, "Between Instincts and Reason: Understanding a Critical Relationship," *Academia Letters*, accessed June 3, 2023, https://www.academia.edu/51621276/Between_instincts_and_reason_understanding_a_critical_relationship.

[29] Richard Dawkins, *The Extended Phenotype: The Long Reach of the Gene*, Revised edition (Oxford; New York: Oxford University Press, 1999).

[30] Louise Barrett, *Beyond the Brain. How Body and Environment Shape Animal and Human Minds* (Princeton–New York: Princeton University Press, 2011), 29–31; "[It's e]nvironments [that] are inherited – a notion that shakes the nature-nurture dichotomy to its core" (Blumberg, "Development Evolving," 7).

[31] Due to space constraints, we must bypass other philosophical aspects of the FWCs' *automatons* premise. We find that pointing out the opposition to the whole Cartesian approach as a cornerstone of pragmatism should be sufficient. Perhaps the best (unfriendly) summary of the pragmatic anti-Cartesian stance can be found in Jerry Fodor, see Fodor, *Lot 2*, 11–18.

[32] Kenneth Davis and Jaak Panksepp, *The Emotional Foundations of Personality* (New York–London: Norton & Company, 2018), 9.

of Animal Behavior: "Under carefully controlled experimental circumstances, the animal behaves as it damned well pleases."[33]

With that said, we need to note that lately, it is becoming increasingly clear that the behavioural paradigm, while not inherently flawed, can offer adequate accounts for only low-level learning processes ubiquitous in the Animal Kingdom, whilst it falls short for the more complex ones. Of course (but leaving behaviourism aside), there exist pretty successful associative attempts to interpret complex cognitive mechanisms found in animals, including transitive inference, episodic memory, causal learning, goal-directed behaviour, imitation, and even metacognition or essential learning of words,[34] but herein lies their fundamental weakness in the context we are interested in; if the same associative models accurately describe higher forms of cognition in both humans and other animals, where is the difference? Even worse – the exact mechanisms explained by associationists within their paradigm conceptual schema can be conveniently captured in terms of the Wagner-Rescorla model as cognitive (representational) without losing the explanatory grip of the purely associative approach. As a result, the vision of animal machines as operating divergently from us fades.[35]

In the 1970s, Robert Rescorla and Allan Wagner, through analysis of experiments on rats, adopted a more nuanced approach to behavioural explanations. They posited a hypothesis (empirically validated since) that learning occurs only when the brain detects discrepancies (surprisal) between predicted events and the following experiences. The error signal, indicative of surprise, plays a decisive role in this process; organisms can learn only when the events experienced deviate from their expectations.[36] The upshot is that the brain is not an organ that passively associates closely temporally-related stimuli, even in the classical conditioning paradigm.

As it turns out, the Rescorla-Wagner rule, psychological in design, neatly fits modern neural network models. According to those theories, the brain actively uses sensory input to predict the probability of forthcoming stimuli. This involves computing the synaptic weights across neural layers and assessing the difference between anticipated (top–down prediction) and actual sensory data (bottom–up evidence). The prediction

[33] Philip Ball, *The Book of Minds: Understanding Ourselves and Other Beings, from Animals to Aliens* (London: Picador, 2023), 425.

[34] See Cameron Buckner, "Understanding Associative and Cognitive Explanations in Comparative Psychology," in *The Routledge Handbook of Philosophy of the Animal Mind*, eds. Kristin Andrews and Jacob Beck (London–New York: Routledge, 2017), 409–18.

[35] Aaron Blaisdell, "Comparative Approaches to Study of Basic Processes of Cognition: A Tale of Three Species," in *Animal Cognition: Principles, Evolution, and Development*, ed. Mary Olmstead (New York: Nova Science Publishers, Inc., 2016), 27–59.

[36] Robert Rescorla and Allan Wagner, "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement," in *Classical Conditioning II: Current Research and Theory*, eds. Artur Black and William Prokasy (New York: Appleton-Century-Crofts, 1972), 64–99.

error signal (information of mismatch) acts as a gauge of how surprised the brain is by the experienced difference, leading to adjustments of synaptic weights, which in turn causes corrections to the internal representation of the stimulus in direct proportion to both the magnitude and value of the prediction error. In effect, this mechanism allows the subsequent prediction to gain enhanced accuracy in ensuing analogous contexts. Expression of surprise in rats, suggested decades earlier by Edward Tolman and staunchly criticised by his contemporaries,[37] turned out to be one of the primary manifestations of the operations of neural mechanisms driving learning processes.[38] We will refer to this thread in the following articles.

Meantime, in contemporary comparative psychology, strenuous efforts to cling to associationism (non-mentalistic explanations) have led to situations where explanations of experimental outcomes become so convoluted that they are practically indistinguishable from descriptions in terms of some sort of rational behaviour.[39] Indeed, "there is substantial, if not incontrovertible, evidence for [at least] instrumental reasoning in a range of nonhuman animals, especially rodents, corvids, and primates."[40] Additional reservations regarding associationism arise from the reflection that evolution would waste an incredible amount of brain tissue if the exact set of non-cognitive mechanisms were to be responsible for diverse behaviours across the animal spectrum (with humans as the noble exception).[41] The stimulus-response or unconscious associationistic model would now have to account for multiple phenomena – from the learning abilities of fruit flies[42] and responsiveness to the conditioning of rats bereft of the brain cortex[43] through the execution

[37] David Carroll, *Purpose and Cognition: Edward Tolman and the Transformation of American Psychology* (New York: Cambridge University Press, 2017).

[38] See Stanislas Dehaene, *Jak się uczymy? Dlaczego mózgi uczą się lepiej niż komputery*, trans. Dariusz Rossowski (Kraków: Copernicus Center Press, 2021), 287–89.

[39] Instrumental rationality is the concept according to which an individual makes decisions and acts in a way that most effectively leads to achieving specific goals, or to put it more generally, it is the cognitive capacity to select effective means for achieving desired ends given computational and environmental constraints.

[40] Elisabeth Camp and Eli Shupe, "Instrumental Reasoning in Nonhuman Animals," in *The Routledge Handbook of Philosophy of Animal Minds*, eds. Kristin Andrews and Jacob Beck (London–New York: Routledge, 2017), 105.

[41] The associative explanatory strategy describes animal behaviour as produced solely by tracking a fixed set of relations between stimuli or as shaped by reactions to stimuli presented to her in the past (present in her learning history).

[42] Robert Murphey, "Instrumental Conditioning of the Fruit Fly, Drosophila Melanogaster," *Animal Behaviour*, 15 (1967): 153–61, accessed June 10, 2023, https://doi.org/10.1016/S0003-3472(67)80027-7.

[43] James Grau, "Learning from the Spinal Cord: How the Study of Spinal Cord Plasticity Informs Our View of Learning," *Neurobiology of Learning and Memory*, 108 (February 2014): 155–71, accessed June 10, 2023, https://doi.org/10.1016/j.nlm.2013.08.003.

of sophisticated skills of "lower" animals such as goal-oriented, independent from extrinsic-rewards manoeuvrings with tiny electric vehicles,[44] and perspective-taking during play in "hide-and-seek"[45]– capacities that rodents which do possess this organ are happy to master.

In sum, regardless of which traditional strategy for explaining animal behaviour upholders of the approach cited in FWC choose, it will, at least sometimes, be too narrow and incomplete for certain cases. But suppose one suggests that some animal skills are similar to ours and that this resemblance could have the same rooting. In that case, one should be aware that, inevitably, she will be accused of committing an anthropomorphic fallacy. As the late Frans de Wall insightfully noted:

> To say that an animal follows its "instincts" is as much a matter of interpretation as to say that it acts "intentionally," yet it is only the second kind of description that gets one into trouble. Given that the absence of intentionally is as hard to prove as its presence, and given the lack of evidence that animals differ from people in this regard, such caution would be acceptable if human behavior were held to the same standard. But, of course, it is not. Cries of anthropomorphism are heard, particularly when a ray of light hits species other than our own.[46]

We should note that anthropomorphic dispositions are based on social learning mechanisms and belong to natural ways in which people explain the nonphysical behaviour of objects. It is often unwarranted – that is a fact – but responsibility for injecting overly anti-anthropomorphic views into Western culture goes, at least in part, to philosophers. On the other hand, the credit for the noticeable retreat from anti-anthropomorphic positions in comparative psychology in recent decades does not go to the philosophers who came to conclusions that differ from those previously announced and, for example, admitted that animals think. Its sources lie in the works of scientists associated with the cognitive revolution, such as Donald

[44] Rats are surprisingly fast learners if they were previously group-housed in an affordance-rich habitat, see Elizabeth Crawford et al., "Enriched Environment Exposure Accelerates Rodent Driving Skills," *Behavioural Brain Research*, 378 (27 January 2020): 112309, accessed June 10, 2023, https://doi.org/10.1016/j.bbr.2019.112309. We will revisit this particular example in more detail in our third article.

[45] It is worth underscoring that the performance of rats in the latter experiments has been rewarded with social interaction (tickling), and the chosen experimental procedure, which ultimately led to the successful demonstration of their sophisticated mental capacities, contradicts basic behavioural tenets. (Annika Stefanie Reinhold et al., "Behavioral and Neural Correlates of Hide-and-Seek in Rats," *Science*, 365, no. 6458 (13 September 2019): 1180–83, accessed June 10, 2023, https://doi.org/10.1126/science.aax4705.

[46] Frans de Waal, "Anthropomorphism and Anthropodenial: Consistency in Our Thinking about Humans and Other Animals," *Philosophical Topics*, 27, no. 1 (1999): 256, accessed June 10, 2023, https://www.jstor.org/stable/43154308.

Griffin, who drew bold conjectures about the mental abilities of animals by observing the scientific changes that resulted in the behavioural paradigm's degeneration in the 1970s. Griffin called for holding animal subjects to similar methodological standards as those applied to humans – starting from admitting that they are conscious.[47] Regrettably, awareness of the perception change has not reached many philosophers yet (but with a growing number of notable exceptions).

The reluctance to admit that we share, even to the slightest degree, the ability to reason with other animals is often associated with an internal conflict between empirical observations and dualistic intuitions about mind, soul and body.[48] This phenomenon has even earned its term – anthropodenial: a failure to accept that humans are animals too. This defensiveness manifests in a pre-emptive rejection of mental characteristics, which we could plausibly share with other species – a significant feature of Western intellectual culture. Philosopher Martha Nussbaum has found anthropodenialism to be both epistemic and moral deformity[49] which – we must add – makes us substantially uncharitable interpreters of *mere brutes* – a habit hard to eradicate.[50] It is still a prevalent phenomenon in philosophy, and as is often

[47] See Donald Redfield Griffin, *The Question of Animal Awareness: Evolutionary Continuity of Mental Experience* (New York: Rockefeller University Press, 1976). We want to thank the Anonymous Reviewer for pointing this.

[48] See for example Paweł Fortuna et al., "The Relationship between Anthropocentric Beliefs and the Moral Status of a Chimpanzee, Humanoid Robot, and Cyborg Person: The Mediating Role of the Assignment of Mind and Soul," *Current Psychology*, 6 November 2023, accessed December 1, 2023, https://doi.org/10.1007/s12144-023-05313-6.

[49] "[H]uman is the only animal that hates its own animality" (Martha Nussbaum, "Compassion: Human and Animal," in *Understanding Moral Sentiments: Darwinian Perspectives?*, eds. Hilary Putnam, Susan Neiman, and Jeffrey Schloss (London–New York: Routledge, 2017), 146.

[50] To clarify, we are not criticising "naive anthropomorphism" here. It is obvious that children and adults who are emotionally attached to their pets make these types of attributions. They also do it in contexts that do not reflect the essence of mental similarity, e.g. in humanising engagement with mobile phones or supernatural entities. Making accurate attribution is, of course, a methodological problem, but we do not deal with this in the article. We argue, however, that in the Western culture, there exists a (weakening) tendency – inherited from the domination of behaviourism and as a result of the influence of the philosophical radical scepticism, to treat anthropomorphism as a form of immaturity or even intellectual limitations if it comes from educated adults. In the article, we criticise scientific and philosophical dogmatic anthropodenialism due to the culture-forming impact of these intellectual habits. We want to thank the Anonymous Reviewer for drawing attention to the need for this clarification. For further discussions on this issue, see; Farshad Nemati, "Anthropomorphism in the Context of Scientific Discovery: Implications for Comparative Cognition," *Foundations of Science*, 28, no. 3 (2023): 927–45, accessed March 29, 2024, https://doi.org/10.1007/s10699-021-09821-1; Janik Festerling and Iram Siraj, "Anthropomorphizing Technology: A Conceptual Review of Anthropomorphism Research and How It Relates to Children's Engagements with Digital Voice Assistants," *Integrative Psychological & Behavioral Science*, 56, no. 3 (2022): 709–38, accessed March 30, 2024, https://doi.org/10.1007/s12124-021-09668-y; Esmeralda G. Urquiza-Haas and Kurt Kotrschal, "The Mind behind Anthropomorphic Thinking: Attribution of Mental States to Other Species," *Animal Behaviour*,

the case with sensitive topics – any finding from empirical inquiry, if unpleasant and unwelcome, faces fierce resistance. In this respect, influential philosophical figures such as Donald Davidson, John McDowell, Robert Brandom, and Peter Carruthers continue to shape our perception of the human-animal dichotomy.[51]

Thus, we arrive at the crux of the issue embedded in Bremer's FWC – the Cartesian view of the body as analogous to clockwork. Naturally, this directs us to the predominant philosophical conviction that natural language is the sole enabler of thought. In its 20th-century analytic rendition, in a streamlined schema, it goes like this: only humans can be considered agents, as we are the only species capable of asking and answering questions about our own reasons for actions – *we think*. Since we rejected the existence of syntax in nonhumans,[52] we are justified to conclude that they do not have *real* language, so they are unable to perform such "reflective acts," *ergo – they do not think*. And if they do not think, they do not operate on concepts, including beliefs and desires – *they lack minds*. "Conceptless" species cannot even perceive – they merely sense or register data and, crucially, *do not act* but passively await stimuli and react to them. Beasts are limited to behaviours.

The dispute about animal languages is ongoing, and there is no place to ponder it further here.[53] However, the problem with Descartes' legacy is more profound than this. Even many anti-Cartesian philosophers and scientists have overlooked the fact that, although substance dualism seems to have been eradicated from mainstream thought, the machine metaphor initiated by his considerations has survived and permeated the natural and cognitive science in various forms. It thrives everywhere, from biology and comparative psychology to neurosciences, at every level of explanation – from the inner workings of a single cell to the processes of an organ-

109 (2015): 167–76, accessed March 30, 2024, https://doi.org/10.1016/j.anbehav.2015.08.011; Nicholas Epley, "A Mind like Mine: The Exceptionally Ordinary Underpinnings of Anthropomorphism," *Journal of the Association for Consumer Research*, 3, no. 4 (2018): 591–98, accessed March 30, 2024, https://doi.org/10.1086/699516; Domenica Bruni, Pietro Perconti, and Alessio Plebe, "Anti-Anthropomorphism and Its Limits," *Frontiers in Psychology*, 9 (2018), accessed March 31, 2024, https://doi.org/10.3389/fpsyg.2018.02205.

[51] Sadly, some of the mentioned philosophers identify themselves as neopragmatists, and some are thought of (often imprecisely) as subscribing to the pragmatist tradition. Their argumentation connects with "the linguistic turn" in analytic philosophy, which hinges on the presumption that to think, one must possess natural language capacities, defined as restrictively as possible. Such tactics usually prevent any debate on animal minds, even though, logically speaking, it begs the question, see Adriana Schetz, "O tak zwanym problemie prostych umysłów," *Diametros*, 30 (2011), 41–60, accessed June 10, 2023, https://doi.org/10.13153/DIAM.30.2011.455.

[52] Even this premise is not entirely accurate, see Klaus Zuberbühler, "Syntax and Compositionality in Animal Communication," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375, no. 1789 (2019): 20190062, accessed September 9, 2024, https://doi.org/10.1098/rstb.2019.0062.

[53] See Arkadiusz Gut, *O relacji między myślą a językiem. Studium krytyczne stanowisk utożsamiających myślenie z językiem* (Lublin: Towarzystwo Naukowe KUL, 2009), 47–66.

ism in its entirety. And if the only ontological difference between an agent and an *automaton submerged in the realm of "things"* rests in language competence, there remains no room for *animal agency*. Without it, no genuine voluntary action in the wild is conceivable.[54] To paraphrase Gilbert Ryle: *there are literally no ghosts in those machines*. With that said, let us glimpse at human agency from the perspective of the FW exponents.

Is There a Pragmatist in the Room?

In Western intellectual culture, the understanding of FW is commonly informed by three prototypical sources: 1) the Augustinian, an attempt to reconcile the presence of evil with the benevolence and omnipotence of God; 2) the Cartesian, to which the author of FWC seems leaning; and 3) the Kantian, an ostensibly agnostic effort to hide Deity behind the equation and replacing Her with the Universal Reason while preserving the rest of the construct. Given these efforts' conceptual sophistication and metaphysical depth, engaging in historical analyses or classic debates emanating from the aporias embedded in them would require ample space, with no promise of a sensible conclusion in view. Instead, we highlight specific issues that arise from the attachment to the concept of FW.

At the outset, let us draw from George Lakoff and Mark Johnson's concise pragmatic reconstruction of the category allegedly enabling it, namely the human mind. According to the mainstream Western philosophical tradition: 1) the mind (reason) operates in two separate modes – theoretical and practical, with only the latter somehow interacting with FW; 2) rational thought is mediated by natural language; it is literal, logical, conscious, disembodied (transcendent), and free from emotion;

[54] Incidentally, if the aforementioned philosophers were consistent, proponents of the natural language argument would eagerly welcome Large Language Models of artificial intelligence (LLMs) into the family of moral beings. On the other hand, "the language uniqueness argument" is gradually losing its gravitas due to advancements in research within cognitive neuroscience and the embodied cognition paradigm regarding the embodied foundations of natural language, resulting in significant shortening of the distance between "us" and "them," see Christopher Preston, "Animality and Morality: Human Reason as an Animal Activity," *Environmental Values*, 11, no. 4 (2002): 427–42, accessed June 10, 2023, https://www.jstor.org/stable/30301901. We should add that from an embodied (enactive) point of view, the same conclusion about LLMs as candidates for our cognitive (and moral) kins would be absurd precisely based on their lack of embodiment in the biological sense; see Anthony Chemero, "LLMs Differ from Human Cognition Because They Are Not Embodied," *Nature Human Behaviour*, 7, no. 11 (2023): 1828–29, accessed December 1, 2023, https://doi.org/10.1038/s41562-023-01723-5.

3) due to these distinguished features of the mind, humans are the only biological organisms capable of carrying out rational (and moral) actions; this is because the rational domain is upheld by God or nested in Universal Reason, reflecting the rational structure of Being; 4) humans recognise "Truths" by discovering universal concepts, which reflect the objective reality and thus are endowed with normative power; concomitantly, the privilege of accessing Universal Reason through conceptual means is the basis of human freedom (autonomy) because the intelligible sphere of the mind is free from the cause-effect relations characteristic of the physical domain. Therefore, rationality, conceived as determining autonomy, constitutes "the essence of men" and guarantees humanity's distinctiveness from the rest of the living.[55]

Due to such framing, the philosophical understanding of FW faces many challenges. Consider the justification for human uniqueness, which is deeply seated in the notion of "the space of reasons." This idea, derived from Aristotle's concept of "rational animals," asserts that only humans can reason before acting. It proclaims normatively that *rational action* is underpinned by *reliable beliefs* (obtained through principles of logic, probabilistic inferences, and so forth) and *desires* (goal-oriented preferences), culminating in selecting the *optimal decision to act* to maximise intended outcomes. Simultaneously, we become morally responsible for the chosen action. There may be perhaps some adequacy in this description, but could it serve as justification for distinguishing us from other animals? We mentioned instrumental rationality earlier, but there are many more definitions of this term in philosophy, and none of them seems conclusive.[56]

The abovementioned claim, even when handled in a liberal fashion, leads to more problems. For example, while granting animals first-order desires and will, Harry Frankfurt suggested that we are more than "wantons" – we can choose our wants, and only possessing these secondary-level "wants" grants FW and moral responsibility to us. However, it leaves crucial questions unanswered, mainly – how did we become beings who value one "want" over another? At what point did we gain the power to construct our value systems independently of all forces acting on us, at minimum, since our conception? This approach runs into Galen Strawson's Basic Argument against self-creation and ultimate moral responsibility by underestimating evolutionary, biological, developmental and environmental factors con-

[55] See George Lakoff and Mark Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (New York: Basic Books, 1999), 552–54.

[56] In the early 1990s, Robert Salomon counted 32 definitions of rationality while not accounting for instrumental, bounded, and embodied variants, see Robert Solomon, "Existentialism, Emotions, and the Cultural Limits of Rationality," *Philosophy East and West*, 42, no. 4 (October 1992): 597–621, accessed June 10, 2023, https://doi.org/10.2307/1399671.

tributing to motivational states.[57] We can undoubtedly do what we want, but is what we want ultimately "up to us?"[58] Or, given the external circumstances, are we not rather beings who appreciate certain values because *that is what persons like us* care for? We may as well dub Frankfurt's claim the "Münchhausen's bootstrap puzzler."

Another problem concerning "the space of reasons" pops up when it is referred to as the basis for justification. FW is often linked to moral responsibility, defined in the Frankfurtian manner as the "ability to do otherwise." So, a person can be held morally responsible if she could have chosen a different motivating reason from the mentioned domain and acted on it or refrained from taking any action. However, for creatures like us, the arrow of time goes in one direction. There is hardly any link between post-factum reflection and the ability to act differently, except for the influence of the experienced consequences of a given action on learning proper conduct in the future (a mechanism similar for all species capable of reinforcement learning). Moreover, from a pragmatic perspective, there is no way to test the claim of "the possibility of doing otherwise." Even comparing similar cases will not do the job. There are no identical situations for comparison, identical self-control capacities or learning histories in real life, and for nonlinearly organised brains like ours, even a tiny experience could result in a dramatically different life trajectory. That would be it for our supposed counterfactual powers of controlling our actions. What about deliberations undertaken before the decision? If one were to understand in Witt-

[57] See Galen Strawson, "O niemożliwości całkowitej odpowiedzialności moralnej," trans. Jacek Jarocki, *Roczniki Filozoficzne*, 65, no. 1 (2017): 109–29, accessed June 10, 2023, https://doi.org/10.18290/rf.2017.65.1-6.

[58] "The statement that a person enjoys freedom of the will means that he is free to want what he wants" (Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy*, 68, no. 1 (1971): 11, accessed June 10, 2023, https://doi.org/10.2307/2024717). The assumption that only humans can choose their wants on some ontologically exceptional grounds, different from those of other animals, is highly dubious. It holds an unrealistic view of *Homo sapiens* and other "higher" animals, given that contemporary neurobiology paints a more Schopenhauerian picture: we can do what we want, but we can wish only what we must; see: Sapolsky, *Determined*. Furthermore, from a pragmatistic stance on this issue, Peirce agreed with the deterministic claim that every act of will is caused by the strongest motivation leading to it. At the same time, he pondered why we believe that every cause of action must necessarily lie within consciousness – a question whose validity has been confirmed to its fullest by modern cognitive science. As Peirce has noticed, every action is an action for a reason; therefore, it is a voluntary action, but the decision to consider one reason as more potent than another can be (and often is) influenced by factors entirely beyond our control or even knowledge. From the logical standpoint, for it is impossible to determine precise rules governing the emergence of specific reasons standing behind a given motivational state, he viewed the claim about acting for a reason based on the strongest motivation as a tautology, see Charles Sanders Peirce, "Dalsze konsekwencje braku czterech zdolności (1868)," in *Charles Sanders Peirce. O Nieskończonej wspólnocie badaczy*, trans. Agnieszka Hensoldt (Opole: Wydawnictwo Uniwersytetu Warszawskiego, 2009), 97–98. That said, we must add that views on the freedom of the will within pragmatism have been and still are diverse; there is no unanimity in this respect.

gensteinian style the nature of the deliberation process preceding any decision, one would have no choice but to agree that we cannot know the results of the operation until our brains finish the calculations. This process gets even more complex – one must consider that we are only aware of (represent) some of the variables included in the equation, as most are being sampled from the outside of the system or arise from within as a response to unconsciously processed environmental cues from the past, predominantly bypassing our awareness. Further, once we see the results, we cannot "unsee" them. There is no "otherwise" in this sense.[59] At best, this "otherwise" notion is a postdictive visceral cue for future behaviour (exhibited in experience in, for example, regret).[60] "The space of reason" comes into play only when we look for presumptive causes of the act performed by us – when a third party asks: "Why did you do that?" – so "the space of reason" becomes "the space of justifications." We could say we are our brain/bodies' interpretation systems of themselves. But before all else, they recognise action ownership, and we interpret it as FW; justifications of our deeds come later. In sum, both cases show that, during reflection on possible alternative ways of acting that we could have taken or when we deliberate about future actions, we doubtfully possess control over why we are the sort of agents who are predisposed to choose X over Y or why we have chosen X over Y.

Here comes another FW problem – it is an undeniable fact that we spend plenty of time reasoning with others. But what is happening during those processes? If

[59] From the perspective of Peirce's pragmatic maxim (discussed below), the meaning of FW refers precisely to the concept of reflection (retrospective guessing) after an action has been performed, see Charles Peirce, "Illustrations of the Logic of Science II – How to Make Our Ideas Clear," in *Popular Science Monthly*, 12, 1878, accessed June 15, 2023, https://en.wikisource.org/wiki/Popular_Science_Monthly/Volume_12/January_1878/Illustrations_of_the_Logic_of_Science_II. If I am ashamed of something I have done, I have to ask myself if I could have done it differently, whether I could have resisted the temptation to do what I just did, although the question can take an implicit form in most cases. Therefore, if we expose what is most important to us in the case in question, there appears a temptation shared by the philosophical tradition (with which Peirce seemed to disagree) to assume that I could have acted differently because I acted wrong (contrary to others' and own expectations) – hence feeling of shame. However, when the same issue is sorted from a perspective devoid of subjectivity and referring to external facts instead, it becomes evident that it was impossible to resist a sufficiently strong temptation, which does not mean that the experience of shame will not increase willpower in the future. Therefore, the experience of regret due to suboptimal actions (or retrospective visceral counterfactual prediction) plays an important role, especially in shaping *self-control*, as the ability to look at the world from a broader perspective. According to Peirce, *self-control is the only form of freedom worthy of this term*, see Peirce, *Dalsze konsekwencje*, 98. Crucially, the capacity for exerting self-control is not an exclusively human property, see Beran, *Self-control in Animals*.

[60] The capacity for regret – an affective state intertwined with failure in exerting efficient self-control, which is something akin to Dewey's notion of "secondary experience" – has also been discovered in rodents, see Adam Steiner and David Redish, "Behavioral and Neurophysiological Correlates of Regret in Rat Decision-Making on a Neuroeconomic Task," *Nature Neuroscience*, 17, no. 7 (July 2014): 995–1002, accessed June 11, 2023, https://doi.org/10.1038/nn.3740.

the dual process theories are on the right track, our reasoning capacities come in two metaphorical flavours based on one mechanism. These are System 1, which is fast and effortless (heuristics, intuitions), and System 2, which is slow and effortful.[61] We rarely initiate the latter – only when, in a given context, there is a complex and pivotal problem to solve. The question is, therefore, if FW exists, then when we use System 1 (which is most of the time), are we free or not? If we are still "free-willers," we must recognise that animals also use heuristics.[62] So maybe we are free only when using System 2? The best evidence for this would be to demonstrate that people using it more often than others make better decisions in their personal lives, for example, in the moral domain. And here comes the disappointment – research on moral philosophers (these are paradigmatic System 2-reasoners most of the time while living off of it) shows no difference in moral behaviour between them and their peers.[63] And remember – these are the autochthons of "the space of reason."

Maybe it is because, during reasoning, moral philosophers can spot the "correct reasons" for action, yet they cannot enact them. This would mean they sometimes act contrary to the "correct reasons" and sometimes on the "reasons" they disapprove of. There must be some additional process at play by which moral philosophers not only recognise "the correct reason" but also get the impetus to act on it, regardless of its "correctness." But can they revolt against such "reason" in such cases? Even Frankfurt conceded that sometimes "reasons" hinder us from making alternative choices when he mentioned Martin Luther, accused of heresy before the Diet of Worms, stating: "Here I stand. I can do no other." Although Frankfurt did not fully explore the implications of his remark.

[61] Daniel Kahneman, *Pułapki myślenia. O myśleniu szybkim i wolnym*, trans. Piotr Szymczak (Poznań: Media Rodzina, 2012).

[62] Gerd Gigerenzer and Wayne D. Gray, "A Simple Heuristic Successfully Used by Humans, Animals, and Machines: The Story of the RAF and Luftwaffe, Hawks and Ducks, Dogs and Frisbees, Baseball Outfielders and Sidewinder Missiles – Oh My!," *Topics in Cognitive Science*, 9, no. 2 (2017): 260–63, accessed March 28, 2024, https://doi.org/10.1111/tops.12269.

[63] See Schwitzgebel, "How Often Do Ethics Professors Call Their Mothers?," *Aeon*, accessed 14 June, 2023, https://aeon.co/essays/how-often-do-ethics-professors-call-their-mothers; Eric Schwitzgebel, "Do Ethicists Steal More Books?," *Philosophical Psychology*, 22, no. 6 (2009): 711–25, accessed April 1, 2024, https://doi.org/10.1080/09515080903409952; Eric Schwitzgebel, "The Moral Behavior of Ethicists and the Role of the Philosopher," in *Experimental Ethics: Toward an Empirical Moral Philosophy*, eds. Christoph Luetge, Hannes Rusch, and Matthias Uhl (London: Palgrave Macmillan UK, 2014), 59–64, Eric Schwitzgebel and Joshua Rust, "The Moral Behavior of Ethics Professors: Relationships among Self-Reported Behavior, Expressed Normative Attitude, and Directly Observed Behavior," *Philosophical Psychology*, 27, no. 3 (2014): 293–327, accessed April 1, 2024, https://doi.org/10.1080/09515089.2012.727135.

Zbigniew Słuszkiewicz

Free Will and the Pragmatic Maxim

In this century, cognitive sciences have brought to bear an amass of counter-evidence that should prompt us to question whether we, in fact, respond to reasons under this classical description, its metaphysical sophistication notwithstanding. More substantiated explanations highlight our remarkable ability to concoct *post hoc* rationalisations for our actions – an outcome of our cognitive structure, our lack of direct access to these processes themselves, and our immersion in storytelling practices.[64] The multifaceted phenomenon is equally prevalent among laypeople, scientists, and philosophers alike.[65] From this standpoint, not only "the space of reasons" but also FW itself appears to be a story we are inclined to believe. This predicament is not new either; the firmly entrenched rationalistic framework would have already suffered a blow over a half-century ago had its followers recognised the implications of Herbert Simons' discovery that we are not rational utility's maximisers – *Homo oeconomicus* – but rather "satisficers," whose decision-making is limited by constraints of bounded rationality.[66] The above remark gains even more significance in the paper's context if one considers the fact that the concept of bounded rationality is currently taking on an intense, embodied and more species-general flavour.[67]

[64] We tend to rationalise our choices exactly in the same way as capuchin monkeys in the study mentioned in the introduction of the paper (footnote 23), but we use natural language for that, see Richard Nisbett and Timothy Wilson, "Telling More than We Can Know: Verbal Reports on Mental Processes," *Psychological Review*, 84 (1977): 231–59, accessed June 11, 2023, https://doi.org/10.1037/0033-295X.84.3.231; Jonathan Haidt, *Prawy umysł. Dlaczego dobrych ludzi dzieli religia i polityka?*, trans. Agnieszka Nowak-Młynikowska (Sopot: Smak Słowa, 2014). The *post hoc* rationalisations are observed in an even more direct fashion in split-brain patients. In these cases, although a person behaves coherently at first glance, in experimental designs, it has been demonstrated that her left hemisphere fluently confabulates stories to explain actions initiated by the right hemisphere if it has no access to its visual field, see Michael Gazzaniga, *Istota człowieczeństwa. Co sprawia, że jesteśmy wyjątkowi*, trans. Agnieszka Nowak (Sopot: Smak Słowa, 2011).

[65] Eric Schwitzgebel and Jonathan Ellis, "Rationalization in Philosophical and Moral Thought," in *Moral Inferences*, eds. Jean-Francois Bonnefon and Bastien Trémolière (Hove: Psychology Press, 2017), 170–90.

[66] Bounded rationality posits that an agent's decision-making abilities are constrained by available information, the timeframes for making decisions, and cognitive limitations in information processing. These factors lead to the pursuit of merely "satisfactory" solutions; see Herbert Alexander Simon, *Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in a Social Setting*, 1st Edition (New York: Wiley and Sons, 1957). Note that given Simon's approach to his field of interest, it could be equally well argued that he is a classical pragmatist; see Thomas Nickles, "Bounded Rationality, Scissors, Crowbars, and Pragmatism: Reflections on Herbert Simon," *Mind & Society*, 17, no. 1 (2018): 85–96, accessed June 11, 2023, https://doi.org/10.1007/s11299-019-00206-3.

[67] Enrico Petracca, "Embodying Bounded Rationality: From Embodied Bounded Rationality to Embodied Rationality," *Frontiers in Psychology*, 12 (2021), accessed June 11, 2023, 710607, https://doi.

From a broadly taken contemporary cognitive sciences standpoint, a picture of us emerges as narrative beings operating *via* mental shortcuts and heuristics.[68] Paradoxically, even as we indulge in grandiose stories about our cognitive uniqueness, we stumble *en passant* upon, for example, the capabilities of rats to counterfactual reasoning[69] or their susceptibility to committing similar logical fallacies in probabilistic inferences (heuristic thinking mentioned before).[70] The mentioned discoveries strongly suggest that cognisers of various species may also be constrained by norms of bounded rationality.[71]

In this context, both metaphysical and folk beliefs about FW emerging from philosophical or theological considerations appear incompatible with our current understanding of our cognitive capacities or seem, at best, superfluous, simultaneously casting doubt on the traditionally assumed veracity of ethics, which may instead be viewed as a form of *pious* wishful mantra. *Homo sapiens* might not be as much "rational animals" but rather "fiction-making, fiction-believing beings."[72]

org/:10.3389/fpsyg.2021.710607; Samuel Nordli and Peter Todd, "Embodied and Embedded Ecological Rationality: A Common Vertebrate Mechanism for Action Selection Underlies Cognition Heuristic Decision-Making in Humans," in *Embodied Bounded Rationality*, eds. Shaun Gallagher, Riccardo Viale, and Vittorio Gallese (Lausanne: Frontiers Media SA, 2023), 133–43; Riccardo Viale, Shaun Gallagher, and Vittorio Gallese, "Bounded Rationality, Enactive Problem Solving, and the Neuroscience of Social Interaction," *Frontiers in Psychology*, 14 (2023), accessed June 11, 2023, https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1152866.

[68] For a slightly overly optimistic view, see Kahneman, *Pułapki myślenia*.

[69] Vincent Laurent and Bernard Balleine, "Factual and Counterfactual Action-Outcome Mappings Control Choice between Goal-Directed Actions in Rats," *Current Biology*, 25, no. 8 (2015): 1074–79, accessed June 12, 2023, https://doi.org/10.1016/j.cub.2015.02.044; David Redish, "Vicarious Trial and Error," *Nature Reviews Neuroscience*, 17, no. 3 (2016): 147–59, accessed June 12, 2023, https://doi.org/10.1038/nrn.2015.30; David Redish et al., "Sunk Cost Sensitivity during Change-of-Mind Decisions Is Informed by Both the Spent and Remaining Costs," *Communications Biology*, 5, no. 1 (2022): 1–17, accessed June 12, 2023, https://doi.org/10.1038/s42003-022-04235-6.

[70] Conjunction fallacy is basically a fallacy of heuristic reasoning, claimed hitherto to be unique for humans; see Valeria González et al., "The Conjunction Fallacy in Rats," *Psychonomic Bulletin & Review*, 30, no. 4 (2023): accessed August 12, 2023, 1564–74, https://doi.org/:10.3758/s13423-023-02251-z. Described inferential bias could be related to the ability to track statistical regularities in sensory data (statistical learning), a cognitive feature common for both humans and other animals, see Chiara Santolin and Jenny Saffran, "Constraints on Statistical Learning Across Species," *Trends in Cognitive Sciences*, 22, no. 1 (2018): 52–63, accessed August 12, 2023, https://doi.org/10.1016/j.tics.2017.10.003.

[71] This fact would explain the parochial nature of morality as an evolutionary phenomenon; see Gerd Gigerenzer, "Moral Satisficing: Rethinking Moral Behavior as Bounded Rationality," *Topics in Cognitive Science*, 2, no. 3 (2010): 528–54, accessed June 11, 2023, https://doi.org/10.1111/j.1756-8765.2010.01094.x.

[72] Jose Yong, Norman Li, and Satoshi Kanazawa, "Not so Much Rational but Rationalizing: Humans Evolved as Coherence-Seeking, Fiction-Making Animals," *American Psychologist*, 76 (2021): 781–93, accessed June 13, 2023, https://doi.org/10.1037/amp0000674.

This trait actually makes us unique creatures. At the same time, it does not imply that we or other living beings are inherently irrational. Mainly because, in the case of storytelling, there is a profoundly rational reason for populating our epistemic worldviews with belief-desires-intentions-based stories and self-narratives – these practices enable us to develop more coherent and less uncertain robust models of the world,[73] allowing us to act efficiently in concordance with them, sometimes in ways other species will not.[74] This capacity reflects a form of ecological rationality that facilitates fast and frugal heuristics for tasks typical of our species' *umwelt*.[75] However, if we would equalise prevalent tendencies for rationalisation with a shortcoming of "genuine rationality," then the rare instances of thorough, rational reasoning that traditional philosophers hold so dear should be seen more as "scattered beacons on the irrational coastline of human history."[76]

That said, it is not to suggest that the FW story should be dismissed at this point. There could indeed be empirical evidence supporting the utility of FW for human flourishing, at least in a manner demonstrated by the applications of Buddhist meditations within the context of neuroscience.[77] At this stage, we should adopt an agnostic stance towards the ontological status of FW, assuming, at the very least, its potential adaptive value for societies. This could be an effective strategy for tackling the problem because, from the pragmatic philosophy viewpoint, the concept itself is not paramount. The meaning of it does not reside solely within the concept's content.[78] Rather, it is the ramifications (in the case of FW) of the concept that confer

[73] Though it is rather an inversely engineered explicit version of Dennettian "free-floating rationales," see Daniel Dennett, "The Free Floating Rationales of Evolution," *Rivista Di Filosofia*, 103, no. 2 (2012): 185–200, accessed June 13, 2023, https://doi.org/10.1413/37254. Also: Fiery Veit et al., "The Rationale of Rationalization," *Behavioral and Brain Sciences*, 43 (2019): e28, accessed June 13, 2023, https://doi.org/10.1017/S0140525X19001730.

[74] That does not mean that the capacity of storytelling is a value in itself. We can either achieve great things or do detrimental and self-destructive things depending on the engagement with the specific social construct (world model).

[75] Gerd Gigerenzer, Peter Todd, *Simple Heuristics That Make Us Smart*, Illustrated edition (New York: Oxford University Press, 2000).

[76] Stephen Stich, "Could Man Be an Irrational Animal? Some Notes on the Epistemology of Rationality," *Synthese*, 64, no. 1 (1985): 115, accessed June 14, 2023, https://www.jstor.org/stable/20116149.

[77] Matthieu Ricard and Wolf Singer, *Beyond the Self: Conversations between Buddhism and Neuroscience*, 1st edition (Cambridge, Mass. The MIT Press, 2017).

[78] Pragmatic Maxim: "Consider what effects, which might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object" Charles, "Illustrations of the Logic of Science II – How to Make Our Ideas Clear," in *Popular Science Monthly*, 12, 1878, 132, accessed June 15, 2023, https://en.wikisource.org/wiki/Popular_Science_Monthly/Volume_12/January_1878/Illustrations_of_the_Logic_of_Science_II.

meaning, as they serve to determine it.[79] As Peirce wrote: "Our idea of anything is our idea of its sensible effects."[80]

Hence, instead of clinging to its empirically unverifiable features, we should investigate whether belief in FW yields effects for its proponents, society and the world at large – which would serve as the best evidence we can have for its general existence. For example, assuming that belief in FW influences people's behaviours and attitudes, widespread practices of undermining it may have far-reaching implications for social cohesion. Daniel Dennett[81] indeed expressed concern that arguing against FW might be inherently irresponsible and could precipitate catastrophic outcomes.[82] The late renowned philosopher grounded his caution precisely in research results designed to address this issue.

Sharing Dennett's worry over whether belief or disbelief in FW (FWB) influences the believer's behaviour, psychologists initiated investigations a few years back – a pursuit that continues today.[83] Experimental methods in this area typically rely on manipulating the intensity of FWB, a paradigm introduced by Kathelyn Vohs and Jonathan Schooler in 2008.[84] In these experiments, laypeople's concepts are assessed and then altered (either weakened or strengthened) through the presentation of pro and anti-FW content. Subsequently, subjects are exposed to opportunities for misconduct and compared to a control group. Initially, some studies have indicated that FWB can result in negative downstream consequences. In certain cases, increased unethical behaviours, such as cheating, and antisocial attitudes, including racism and aggressiveness towards others, have been observed. The erosion of FWB appears to diminish prosocial attitudes, manifested in altruistic and cooperative behaviours.[85] Some cognitive-neuroscientific studies suggest that such behavioural changes might be linked to self-control-related neural and cognitive mechanisms,

[79] In analogy to the disputes over the doctrine of transubstantiation – if a proposition has no experiential sensible effects, then, for the pragmatist, it is meaningless.

[80] Charles Peirce, *How to Make Our Ideas Clear*, 131.

[81] Daniel Dennett, *Dźwignie wyobraźni i inne narzędzia do myślenia*, trans. Łukasz Kurek (Kraków: Copernicus Center Press, 2015), 479–481.

[82] Leading researchers in this field even raised an alarming warning "that newfound free will-scepticism may end up threatening the humanitarian revolution, potentially culminating in anarchy" Azim Shariff and Kathleen Vohs, "The World Without Free Will," *Scientific American*, 310, no. 6 (2014): 79, accessed June 14, 2023, https://doi.org/10.1038/scientificamerican0614-76.

[83] Different variations of the concept of FW are usually defined psychologically as a class of laypeople's metacognitive judgments about how they intentionally guide their thoughts and actions. In this sense, their belief reflects a perception that people are responsible for actions they can control.

[84] Dennett is explicitly referring to this paradigm.

[85] For an overview, see Esthelle Ewusi-Boisvert and Eric Racine, "A Critical Review of Methodologies and Results in Recent Research on Belief in Free Will," *Neuroethics*, 11 (2018): 97–110, accessed June 15, 2023, https://doi.org/10.1007/s12152-017-9346-3.

for example, by reducing error-detection signals.[86] Conversely, firm FWBs are thought to be associated with intense retributive attitudes, with a higher propensity to victim-blaming and correlate significantly with belief in the non-physical mind.[87]

Be that as it may, confidence in FW seems to exert wide-ranging effects. But yet, subsequent scientific probes into those paradigms have struggled to replicate their original landmark findings.[88] To add insult to injury, an up-to-date, extensive meta-analysis encompassing 145 published and unpublished experiments, while providing clear evidence that FWBs are susceptible to manipulation, has simultaneously found little-to-no evidence that such manipulations result in short-term downstream consequences, let alone long-term changes in behavioural patterns, attitudes, or cognition of the participants.[89] Repeating failures to reproduce experimental results may have multiple causes, and the jury is still out. Nevertheless, at this stage, it may also lead to the conclusion that the failings of psychological research on FWB vividly illustrate the ongoing replication crisis in social psychology.

[86] Davide Rigoni, Gilles Pourtois, and Marcel Brass, "'Why Should I Care?' Challenging Free Will Attenuates Neural Reaction to Errors," *Social Cognitive and Affective Neuroscience*, 10, no. 2 (2015): 262–68, accessed June 15, 2023, https://doi.org/10.1093/scan/nsu068.

[87] Cory Clark et al., "Free to Punish: A Motivated Account of Free Will Belief," *Journal of Personality and Social Psychology*, 106, no. 4 (2014): 501–13, accessed June 15, 2023, https://doi.org/10.1037/a0035880; Oliver Genschow and Benjamin Vehlow, "Free to Blame? Belief in Free Will Is Related to Victim Blaming," *Consciousness and Cognition*, 88 (2021): 103074, accessed June 15, 2023, https://doi.org/10.1016/j.concog.2020.103074; David Wisniewski, Robert Deutschländer, and John-Dylan Haynes, "Free Will Beliefs Are Better Predicted by Dualism than Determinism Beliefs across Different Cultures," ed. Jonathan Jong, *PLOS ONE* 14, no. 9 (2019): e0221617, accessed June 15, 2023, https://doi.org/10.1371/journal.pone.0221617.

[88] Andrew Monroe, Garrett Brady, and Bertram Malle, "This Isn't the Free Will Worth Looking for: General Free Will Beliefs Do Not Influence Moral Judgments, Agent-Specific Choice Ascriptions Do," *Social Psychological and Personality Science*, 8, no. 2 (2017): 191–99, accessed June 15, 2023, https://doi.org/10.1177/1948550616667616; Damien Crone and Neil Levy, "Are Free Will Believers Nicer People? (Four Studies Suggest Not)," *Social Psychological and Personality Science*, 10 (2019): 612–19, accessed June 15, 2023, https://doi.org/10.1177/1948550618780732; Thomas Nadelhoffer et al., "Does Encouraging a Belief in Determinism Increase Cheating? Reconsidering the Value of Believing in Free Will," *Cognition*, 203 (2020): 104342, accessed June 16, 2023, https://doi.org/10.1016/j.cognition.2020.104342; Charlotte Eben et al., "Are Post-Error Adjustments Influenced by Beliefs in Free Will? A Failure to Replicate Rigoni, Wilquin, Brass and Burle, 2013," *Royal Society Open Science*, 7, no. 11 (2020): 200664, accessed June 16, 2023, https://doi.org/10.1098/rsos.200664; Maayan Katzir and Oliver Genschow, "Automatic or Controlled: How Does Disbelief in Free Will Influence Cognitive Functioning?," *British Journal of Psychology*, 113, no. 4 (2022): 1121–42, accessed June 16, 2023, https://doi.org/10.1111/bjop.12578.

[89] Oliver Genschow et al., "Manipulating Belief in Free Will and Its Downstream Consequences: A Meta-Analysis," *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc,* 27, no. 1 (2023): 52–82, accessed June 17, 2023, https://doi.org/10.1177/10888683221087527.

On top of that, the entire approach to investigating folk belief in FW might contain a more fundamental flaw. Experimental philosophy (X-phi) and social psychology have been operating so far under the prevalent presumption that belief in FW is a universal phenomenon. However, more extensive and nuanced cross-linguistic studies have demonstrated that not only about 91% of participants in these experiments were WEIRD (Western, Educated, Industrialised, Rich, and Democratic) and thus did not represent the diversity of the global population. Above all, the lexical expressions of FW in Chinese, Hindi, and Mongolian do not correspond to the same concept as its Western counterparts. Instead, analogous terms in these languages appear to be superficial translations from the Western tradition, which relates to individualistic properties that natives in collectivist cultures typically do not engage with.[90] These results suggest that the concept of FW, as commonly defined, is a culturally-specific construct rather than a psychologically universal one, or to put it directly – it is a conceptual invention rather than an empirical discovery.[91]

Instead of Conclusions

If, in fact, FW is not a universal phenomenon, cannot be adequately explained by our rational capabilities, cannot be tested in "otherwise" cases, and may not contribute to our everyday conduct, then is it an empty narrative? An illusion? Does it mean we have no FW and should get rid of this notion? Our answer is both yes and no. On the one hand, maintaining that the decision to act follows from the ontology of FW – a byproduct of folk intuitions merged with theological supposition – is like investing in Noumenland. Do not expect a payoff. To return to the marine

[90] Renatas Berniūnas et al., "The Weirdness of Belief in Free Will," *Consciousness and Cognition*, 87 (2021): 103054, accessed June 16, 2023, https://doi.org/10.1016/j.concog.2020.103054. One could argue that there is compelling evidence on the contrary, as Hagop Sakissian and colleagues examined FW intuitions in a parallel study among participants from the USA, India, Hong Kong and Columbia and found similarities, see Hagop Sarkissian et al., "Is Belief in Free Will a Cultural Universal?," *Mind & Language*, 25, no. 3 (2010): 346–58, accessed June 16, 2023, https://doi.org/10.1111/j.1468-0017.2010.01393.x. The problem is that all participants were students of West-oriented Universities, not to mention post-colonial influences in regions under study. The assumption that a person would be representative of her culture solely based on being a citizen of a non-western country is one of the sources of confusion.

[91] A similar claim could be made about assertions on the universality of certain Western moral intuitions; see Renatas Berniūnas, "Mongolian Yos Surtakhuun and WEIRD "Morality," *Journal of Cultural Cognitive Science*, 4, no. 1 (2020): 59–71, accessed June 16, 2023, https://doi.org/10.1007/s41809-019-00045-1.

metaphor: due to its speculative anchoring, the FW-centered philosophical expedition has been preordained to flounder in search of empirical support. To make its prospects even less promising, at least since Kant, this ship has had no other course except for sailing around its noumenal attractor. In the process, on the surface, it placed a kilwater before the boat and, at some point, began to follow its own wake. Yet, the foamy water will always trail behind. Recognising this fact allows us to concede that the phenomenon leaves an impression, while preventing us from claiming to have discovered that it propels our vessel.

Yet, on the other hand, even if FW does nothing psychologically, it is a concept firmly entrenched in Western societies' structure, and it appears to be a real experience on the subjective level as well. Its *gravitas* rests on the fact that even if we were to abandon belief in it (as some people did), we could not perpetually entertain the idea that our FW is an illusion (as those people cannot).[92] As Stephen Pinker pointed out – the mere fact that while living in Brooklyn, one has learned that the universe is expanding cannot be an existential cause for her to despair because, for creatures like us, Brooklyn is non-stretchable.[93] Similarly, we cannot continuously bear in mind the empirical fact that we are composed of particles essentially consisting of empty space and attempt to act on that knowledge (e.g., by trying to pass through concrete walls instead of using doors).

It seems that we behave as if we were in control, regardless of our thoughts about the concept itself. Our conduct is based on the imperative to act, powered by the sense of control – a bodily sensation plausibly crucial from an evolutionary standpoint, much like our fundamental moral intuitions.[94] Unless we suffer from some neurochemical imbalance or neurostructural deficits, we will not stop acting, making decisions, forming policies, blaming others, caring for close ones, and experiencing affects no matter what story we believe in. All these instances of social cognition are possible to a degree in principle for (at minimum) warm-blooded species due to a neuronally implemented, dopamine-mediated and future-oriented network responsible for reinforcement learning called the reward

[92] Compatibilist commitments to FWB among philosophers and scientists are indeed widespread, although the list of FW-non-believers in modern times, starting from Albert Einstein, is not short either, and evidence about them being failed human beings is nonexistent. For philosophical debate between opposite views on this issue, see Daniel Dennett and Gregg Caruso, *Just Deserts: Debating Free Will* (Polity Press: Medford, 2021).

[93] Steven Pinker, *Tabula rasa: spory o naturę ludzką*, trans. Agnieszka Nowak (Warszawa: Gdańskie Wydawnictwo Psychologiczne, 2012), 275–276.

[94] Jeffrey Schloss, "Darwinian Explanations of Morality: Accounting for the Normal but Not Normative," in *Understanding Moral Sentiments: Darwinian Perspectives?* eds. Hilary Putnam, Susan Neiman, and Jeffrey Schloss (London: Routledge, 2014), 81–121; Michael Ruse, "Common Sense Morality and Its Evolutionary Underpinnings," in *Scientific Challenges to Common Sense Philosophy*, eds. Rik Peels, Jeroen de Ridder, and Rene van Woudenberg (London–New York: Routledge, 2020), 160–83.

system.[95] To express it more holistically – social learning engages the whole organism-environment unit based on the agent's tightly intertwined subcortical and cortical networks.[96] These are highly sensitive to external dynamics, in exchanges with which they form the bedrock for normativity in complex creatures.[97] From that standpoint, blaming, to some degree, but crucially praising, plays an essential role in shaping the conduct of society members.

By saying that there is no need to insist on an explicit rejection of the belief in FW at the individual level, we nevertheless hope that the fear of losing it will share the fate of Fyodor Dostoyevsky's worry that "if there is no God, then anything is permitted." The often-quoted prediction that individual abandonment of religious faith will inevitably lead to the moral decline of societies has not only been falsified, as empirical research clearly shows – the outcomes are quite the opposite.[98] Relinquishing dogmatic adherence to ancient rules frees up moral development processes in countries where religion has lost its dominant position. Apologists for grounding morality in a transcendent domain still tend to rescue their convictions about morals, sermonising its (at least) autonomy with invocations to Hume's Guillotine and Moore's semantic objection to naturalism, known as the "naturalistic fallacy." Both conceptual blockades, however, have long failed to fulfil the hopes placed in them, especially in the face of pragmatist positions.

At this juncture, we assert that belief in FW, or lack of belief, is inconsequential for the individual subject, as we adhere to John Green's foundational (for pragmatism) definition of belief, meaning "that upon which we are prepared to act."[99] And if nothing changes for us, whether we believe in FW or not, then we are not acting upon our stance on this issue. What matters pragmatically is what society, as collective intelligence,[100] does with the concepts shaping our institutions and cultural practices. Therefore, the strategy we suggest regarding the concept of FW is not official rejection, nor is it clinging to its defense – the latter being a practice that

[95] Patricia Churchland, *Conscience: The Origins of Moral Intuition*, 1st edition (New York: W. W. Norton & Company, 2019), 70–95.

[96] Mark Miller and Andy Clark, "Happily Entangled: Prediction, Emotion, and the Embodied Mind," *Synthese*, 195, no. 6 (2018): accessed June 18, 2023, 2559–75, https://doi.org/10.1007/s11229-017-1399-7.

[97] Michał Piekarski, "Motivation, Counterfactual Predictions and Constraints: Normativity of Predictive Mechanisms," *Synthese*, 200, no. 5 (2022): 352, accessed June 18, 2023, https://doi.org/10.1007/s11229-022-03837-1.

[98] Phil Zuckerman, *What It Means to Be Moral: Why Religion Is Not Necessary for Living an Ethical Life* (Berkeley: Counterpoint, 2019).

[99] Cheryl Misak, "The Metaphysical Club," in *The Routledge Companion to Pragmatism*, eds. Scott Aikin and Robert Talisse (New York–London: Routledge, 2023), 8.

[100] We will return to the notion of collective intelligence in the next paper.

seems inevitably approaching *pia fraus*.[101] It is the idea expressed by Dewey in his acknowledgement of the impact evolution theory exerted on the human intellect in *The Influence of Darwinism on Philosophy*, in which he wrote:

> Old ideas give way slowly; for they are more than abstract logical forms and categories. They are habits, predispositions, deeply engrained attitudes of aversion and preference. Moreover, the conviction persists – though history shows it to be a hallucination – that all the questions that the human mind has asked are questions that can be answered in terms of the alternatives that the questions themselves present. But in fact intellectual progress usually occurs through sheer abandonment of questions together with both the alternatives they assume – an abandonment that results from their decreasing vitality and a change of urgent interests. We do not solve them: we get over them. Old questions are solved by disappearing, evaporating, while new questions corresponding to the changed attitude of endeavor and preference take their place.[102]

In conclusion, it is worth recalling that the category of FW is presumed to denote an ability that philosophers and theologians have traditionally denied other animals possess. As we have tried to argue throughout the paper, the Western intuition concerning the ontology of free will becomes less convincing with the passage of time and the development of empirical research. It is becoming increasingly apparent that the metaphysical understanding of this notion emerged due to multifactorial cultural processes. If this is the case, then the consequences of the traditional views on it result from moral luck (or, in the case of other animals, misfortune) rather than from the effective pursuit of truth. Historically, in Western thought, the basis for this concept stems from the considerations of Augustine of Hippo and not the one who originally introduced it. This, in turn, has weighed heavily on the formation of our image of the world and on the treatment of other animals, which has lasted to the present. Perhaps the human-animal dichotomy would take a somewhat less strict form if its initial formulation had prevailed in the process of concept-shaping. The original term "free will" (*libera voluntas*) was put forth for the first time by the Roman poet–philosopher Titus Lucretius, who in turn was inspired by Democritus's atomism and Epicurus's naturalism. His concept of FW was, by design, radically more inclusive to other animals than that of Christian thinkers. As such, it would be a weak frontier post for marking the divide.[103]

[101] Lat. A pious lie.

[102] John Dewey, *The Influence of Darwin on Philosophy and Other Essays in Contemporary Thought* (Scotts Valley: Createspace Independent Pub, 1910/2017), 8–9.

[103] Lucretius was the first Western thinker to explicitly use this term by linking the dynamic character of the universe with the observations of animal behaviour. Simultaneously, he had nothing to say

The last point we are trying to make is that given what has been written so far, the Western notion of FW, as it currently stands, squarely fits the definition of anthropofabulation,[104] a methodological fallacy involving attributing to one's species some extremely cognitively demanding faculty that either does not exist or cannot be proven to exist and then failing to find it elsewhere or outright denying it to animals. To illustrate the point more vividly – asserting that human beings differ from other animals in having free will is like arguing that the difference between a chicken and a horse is such that the latter is, in essence, a Pegasus.

The remark above ties in with the only remaining and hitherto uncontested premise for Bremer's FWC, though insufficient for establishing genuine FW. It relies solely on the subjective experience of authoring a particular action. Yet, it is important to note that the appearance of a perceptual impression is distinct from how things may really be and what the impression signifies for the perceiver. To echo Ludwig Wittgenstein's reaction to Elizabeth Anscombe's supposition that it seemed natural for the ancient to believe that the Sun was orbiting the Earth rather than the Earth turning its axis because it looked like it – "Well, what would it have looked like if it had looked as if the Earth turned on its axis?" – he asked.[105]

In recent years, a naturalistic path has emerged in cognitive science, offering an opportunity to deepen our understanding of the phenomenon in question. It simultaneously carries a promise of bridging the gap between natural processes and the mind. If it is combined with process metaphysics from the philosophy of biology, it could help to uncover the biotic undersoil of "the space of reasons" – a concept that, for millennia, served as a justification for denying agency to other animals. This will be the topic of the two subsequent papers, "Of Rats and Men II: A Pragmatist Reconstruction of the Basis of Agency *via* Free-Energy Principle," and "Of Rats and Men III: A Pragmatist Reconstruction of Advanced Agency *via* the Active Inference."

on the connections of *libera voluntas* with moral responsibility. For him, the animal capacity to act is a straightforward consequence of the universe's structure as described by Epicurus, in which uncaused "swerves" (*clinamen*) of interacting atoms disrupt the causal chain of events. He concluded that this randomness is the basis of an internal motive in all minds, allowing them to exert their effective agency freely by initiating bodily movements as they will.

[104] Cameron Buckner, "Morgan's Canon, Meet Hume's Dictum: Avoiding Anthropofabulation in Cross-Species Comparisons," *Biology & Philosophy*, 28, no. 5 (2013): 853–71, accessed June 18, 2023, https://doi.org/10.1007/s10539-013-9376-0.

[105] Anil Seth, *Being You: A New Science of Consciousness* (London: Penguin Publishing Group, 2021), 79.

Bibliography

Ball, Philip. *The Book of Minds: Understanding Ourselves and Other Beings, from Animals to Aliens*. London: Picador, 2023.

Barrett, Louise. *Beyond the Brain. How Body and Environment Shape Animal and Human Minds*. Princeton: Princeton University Press, 2011.

Benoit, Emile. *Beasts in Eden: The Humane and the Inhumane*. San Diego: Eudaimonia Press, 2016.

Benz-Schwarzburg, Judith. *Cognitive Kin, Moral Strangers? Linking Animal Cognition, Animal Ethics & Animal Welfare*. Leiden: Brill, 2020.

Beran, Michael. *Self-Control in Animals and People*. 1st edition. London: Academic Press, 2018.

Blaisdell, Aaron. "Comparative Approaches to Study of Basic Processes of Cognition: A Tale of Three Species." In *Animal Cognition: Principles, Evolution, and Development*. Edited by Mary C. Olmstead, 27–59. New York: Nova Science Publishers, Inc., 2016.

Blumberg, Mark S. "Development Evolving: The Origins and Meanings of Instinct." *Wiley Interdisciplinary Reviews. Cognitive Science*, vol. 8, no. 1–2 (2017). Accessed July 3, 2023. https://doi.org/10.1002/wcs.1371.

Bremer, Józef. *Czy wolna wola jest wolna?* Kraków: Wydawnictwo WAM, 2013.

Bruni, Domenica, Pietro Perconti, and Alessio Plebe. "Anti-Anthropomorphism and Its Limits." *Frontiers in Psychology*, vol. 9 (2018). Accessed March 31, 2024. https://doi.org/10.3389/fpsyg.2018.02205.

Buckner, Cameron. "Morgan's Canon, Meet Hume's Dictum: Avoiding Anthropofabulation in Cross-Species Comparisons." *Biology & Philosophy*, vol. 28, no. 5 (2013): 853–71. Accessed June 18, 2023. https://doi.org/10.1007/s10539-013-9376-0.

Buckner, Cameron. "Understanding Associative and Cognitive Explanations in Comparative Psychology." In *The Routledge Handbook of Philosophy of the Animal Mind*, edited by Kristin Andrews and Jacob Beck, 409–18. London–New York: Routledge, 2017.

Camp, Elisabeth, and Eli Shupe. "Instrumental Reasoning in Nonhuman Animals." In *The Routledge Handbook of Philosophy of Animal Minds*, edited by Kristin Andrews and Jacob Beck, 1st edition, 100–108. New York: Routledge, 2017.

Carroll, David W. *Purpose and Cognition: Edward Tolman and the Transformation of American Psychology*. New York: Cambridge University Press, 2017.

Chemero, Anthony. "LLMs Differ from Human Cognition Because They Are Not Embodied." *Nature Human Behaviour*, vol. 7, no. 11 (November 2023): 1828–29. Accessed December 1, 2023. https://doi.org/10.1038/s41562-023-01723-5.

Churchland, Patricia S. *Conscience: The Origins of Moral Intuition*. 1st edition. New York: W. W. Norton & Company, 2019.

Clark, Cory J., Jamie B. Luguri, Peter H. Ditto, Joshua Knobe, Azim F. Shariff, and Roy F. Baumeister. "Free to Punish: A Motivated Account of Free Will Belief." *Journal of Personality and Social Psychology*, vol. 106, no. 4 (2014): 501–13. Accessed June 15, 2023. https://doi.org/10.1037/a0035880.

Crawford, Elizabeth, L. E. Knouse, M. Kent, D. Vavra, O. Harding, D. LeServe, N. Fox, et al. "Enriched Environment Exposure Accelerates Rodent Driving Skills." *Behavioural Brain Research*, vol. 378 (2020): 112309. Accessed June 10, 2023. https://doi.org/10.1016/j.bbr.2019.112309.

Crone, Damien L., and Neil L. Levy. "Are Free Will Believers Nicer People? (Four Studies Suggest Not)." *Social Psychological and Personality Science*, vol. 10 (2019): 612–19. Accessed June 15, 2023. https://doi.org/10.1177/1948550618780732.

Davis, Kenneth L., and Jaak Panksepp. *The Emotional Foundations of Personality*. New York–London: W. W. Norton & Company, 2018.

Dawkins, Richard. *The Extended Phenotype: The Long Reach of the Gene*. Revised edition. Oxford–New York: Oxford University Press, 1999.

Dehaene, Stanislas. *Jak się uczymy? Dlaczego mózgi uczą się lepiej niż komputery*, translated by Dariusz Rossowski, Kraków: Copernicus Center Press, 2021.

Dennett, Daniel C. *Dźwignie wyobraźni i inne narzędzia do myślenia*, translated by Łukasz Kurek, Kraków: Copernicus Center Press, 2015.

Dennett, Daniel C. "The Free-floating Rationales of Evolution." *Rivista Di Filosofia*, vol. 103, no. 2 (2012): 185–200. https://doi.org/10.1413/37254.

Dennett, Daniel C., and Gregg D. Caruso. *Just Deserts: Debating Free Will*. Medford, 2021.

Dewey, John. *Human Nature and Conduct*. Overland Park: Digreads.com Publishing, 1922.

Dewey, John. *The Influence of Darwin on Philosophy and Other Essays in Contemporary Thought*. Scotts Valley: Createspace Independent Pub, 2017.

Eben, Charlotte, Zhang Chen, Emiel Cracco, Marcel Brass, Joël Billieux, and Frederick Verbruggen. "Are Post-Error Adjustments Influenced by Beliefs in Free Will? A Failure to Replicate Rigoni, Wilquin, Brass and Burle, 2013." *Royal Society Open Science*, vol. 7, no. 11 (2020): 200664. Accessed June 16, 2023. https://doi.org/10.1098/rsos.200664.

Egan, Louisa C., Paul Bloom, and Laurie R. Santos. "Choice-Induced Preferences in the Absence of Choice: Evidence from a Blind Two-Choice Paradigm with Young Children and Capuchin Monkeys." *Journal of Experimental Social Psychology*, vol. 46, no. 1 (2010): 204–7. Accessed November 3, 2023. https://doi.org/10.1016/j.jesp.2009.08.014.

Epley, Nicholas. "A Mind like Mine: The Exceptionally Ordinary Underpinnings of Anthropomorphism." *Journal of the Association for Consumer Research*, vol. 3, no. 4 (2018): 591–98. Accessed March 30, 2024. https://doi.org/10.1086/699516.

Ewusi-Boisvert, Esthelle, and Eric Racine. "A Critical Review of Methodologies and Results in Recent Research on Belief in Free Will." *Neuroethics*, vol. 11 (2018): 97–110. Accessed June 15, 2023. https://doi.org/10.1007/s12152-017-9346-3.

Fodor, Jerry A. *Lot 2: The Language of Thought Revisited*. Oxford–New York: Clarendon Press, 2008.

Fortuna, Paweł, Zbigniew Wróblewski, Arkadiusz Gut, and Anna Dutkowska. "The Relationship between Anthropocentric Beliefs and the Moral Status of a Chimpanzee, Humanoid Robot, and Cyborg Person: The Mediating Role of the Assignment of Mind and Soul." *Current Psychology* (2023). Accessed December 1, 2023. https://doi.org/10.1007/s12144-023-05313-6.

Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy*, vol. 68, no. 1 (1971): 5–20. Accessed June 10, 2023. https://doi.org/10.2307/2024717.

Frede, Michael. *A Free Will: Origins of the Notion in Ancient Thought*. 1st edition. Berkeley: University of California Press, 2011.

Gazzaniga, Michael. S. *Istota człowieczeństwa. Co sprawia, że jesteśmy wyjątkowi*. Translated by Agnieszka Nowak. Sopot: Smak Słowa, 2011.

Genschow, Oliver, and Benjamin Vehlow. "Free to Blame? Belief in Free Will Is Related to Victim Blaming." *Consciousness and Cognition*, vol. 88 (2021): 103074. Accessed June 15, 2023. https://doi.org/10.1016/j.concog.2020.103074.

Genschow, Oliver, Emiel Cracco, Jana Schneider, John Protzko, David Wisniewski, Marcel Brass, and Jonathan W. Schooler. "Manipulating Belief in Free Will and Its Downstream Consequences: A Meta-Analysis." *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, vol. 27, no. 1 (2023): 52–82. Accessed June 17, 2023. https://doi.org/10.1177/10888683221087527.

Gigerenzer, Gerd, and Wayne D. Gray. "A Simple Heuristic Successfully Used by Humans, Animals, and Machines: The Story of the RAF and Luftwaffe, Hawks and Ducks, Dogs and Frisbees, Baseball Outfielders and Sidewinder Missiles – Oh My!" *Topics in Cognitive Science*, vol. 9, no. 2 (2017): 260–63. Accessed March 28, 2024. https://doi.org/10.1111/tops.12269.

Gigerenzer, Gerd. "Moral Satisficing: Rethinking Moral Behavior as Bounded Rationality." *Topics in Cognitive Science*, vol. 2, no. 3 (2010): 528–54. Accessed June 13, 2023. https://doi.org/10.1111/j.1756-8765.2010.01094.x.

Gigerenzer, Gerd, Peter M. Todd, and ABC Research Group. *Simple Heuristics That Make Us Smart*. Illustrated edition. New York: Oxford University Press, 2000.

González, Valeria V., Sowgol Sadeghi, Linh Tran, and Aaron P. Blaisdell. "The Conjunction Fallacy in Rats." *Psychonomic Bulletin & Review*, vol. 30, no. 4 (2023): 1564–74. Accessed August 12, 2023. https://doi.org/10.3758/s13423-023-02251-z.

Grau, James W. "Learning from the Spinal Cord: How the Study of Spinal Cord Plasticity Informs Our View of Learning." *Neurobiology of Learning and Memory*, vol. 108 (2014): 155–71. Accessed June 10, 2023. https://doi.org/10.1016/j.nlm.2013.08.003.

Griffin, Donald Redfield. *The Question of Animal Awareness: Evolutionary Continuity of Mental Experience*. New York: Rockefeller University Press, 1976.

Gut, Arkadiusz. O *relacji między myślą a językiem*. Lublin: Towarzystwo Naukowe KUL, 2009.

Haidt, Jonathan. *Prawy umysł. Dlaczego dobrych ludzi dzieli religia i polityka?* Translated by Agnieszka Nowak-Młynikowska. Sopot: Smak Słowa, 2014.

Haselager, Pim. "Conceptual Revisions. Intentions and Free Will in the Light of Cognitive Neuroscience." In *Scientific Challenges to Common Sense Philosophy*, edited by Rik Peels, Jeroen de Ridder, and Rene van Woudenberg, 104–20. New York–London: Routledge: Taylor & Francis Group, 2020.

Hume, David. *Traktat o naturze ludzkiej*. Translated by Czasław Znamierowski. Warszawa: ALETHEIA, 2023.

James, William. *Odmiany doświadczenia religijnego. Studium ludzkiej natury*. Translated by Jan Hempel. Warszawa: Aletheia, 2011.

Johnson, Mark, and Tim Rohrer. "We Are Live Creatures: Embodiment, American Pragmatism and the Cognitive Organism." In *Body, Language, and Mind*, vol. 1, edited by Tom Ziemke, Jordan Zlatev, and Roslyn M. Frank, 17–54. Berlin: Mouton de Gruyter, 2007.

Kahneman, Daniel. *Pułapki myślenia. O myśleniu szybkim i wolnym*. Translated by Piotr Szymczak. Poznań: Media Rodzina, 2012.

Katzir, Maayan, and Oliver Genschow. "Automatic or Controlled: How Does Disbelief in Free Will Influence Cognitive Functioning?" *British Journal of Psychology*, vol. 113, no. 4 (2022): 1121–42. Accessed June 16, 2023. https://doi.org/10.1111/bjop.12578.

Keysers, Christian. *Empatia. Jak odkrycie neuronów lustrzanych zmienia nasze rozumienie ludzkiej natury*. Translated by Łukasz Kwiatek. Kraków: Copernicus Center Press, 2020.

Lakoff, George, and Mark Johnson. *Philosophy in the Flesh: The Embodied Challenge to Western Thought*. New York: Basic Books, 1999.

Laurent, Vincent, and Bernard W. Balleine. "Factual and Counterfactual Action-Outcome Mappings Control Choice between Goal-Directed Actions in Rats." *Current Biology*, vol. 25, no. 8 (2015): 1074–79. Accessed June 12, 2023. https://doi.org/10.1016/j.cub.2015.02.044.

Libet, Benjamin, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl. "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness – Potential). The Unconcious Initiation of the Freely Voluntary Act." *Brain*, vol. 106, no. 3 (1983): 623–42. Accessed June 3, 2023. https://doi.org/10.1093/brain/106.3.623.

Llinas, Rodolfo R. *I of the Vortex: From Neurons to Self*. Reprint edition. Cambridge–London: Bradford Books, 2002.

Maffei, Roberto. "Between Instincts and Reason: Understanding a Critical Relationship." *Academia Letters*. Accessed June 3, 2023. https://www.academia.edu/51621276/Between_instincts_and_reason_understanding_a_critical_relationship.

Meyza, Ksenia Z., and Ewelina Knapska. *Neuronal Correlates of Empathy: From Rodent to Human*. Cambridge, Mass: Academic Press, 2018.

Miller, Mark, and Andy Clark. "Happily Entangled: Prediction, Emotion, and the Embodied Mind." *Synthese*, vol. 195, no. 6 (2018): 2559–75. Accessed June 18, 2023. https://doi.org/10.1007/s11229-017-1399-7.

Misak, Cheryl. "The Metaphysical Club." In *The Routledge Companion to Pragmatism*, edited by Scott F. Aikin and Robert B. Talisse, 7–12. New York–London: Routledge, 2023.

Monroe, Andrew E., Garrett L. Brady, and Bertram F. Malle. "This Isn't the Free Will Worth Looking for: General Free Will Beliefs Do Not Influence Moral Judgments, Agent-Specific Choice Ascriptions Do." *Social Psychological and Personality Science*, vol. 8, no. 2 (2017): 191–99. Accessed June 15, 2023. https://doi.org/10.1177/1948550616667616.

Monsó, Susana. "Morality and Mindreading in Nonhuman Animals." *Universidad Nacional de Educación a Distancia (UNED)*, 2016. Accessed June 23, 2023. http://e-spacio.uned.es/fez/eserv/tesisuned:Filosofia-Smonso/MONSO_GIL_Susana_Tesis.pdf.

Murphey, Robert M. "Instrumental Conditioning of the Fruit Fly, Drosophila Melanogaster." *Animal Behaviour*, vol. 15 (1967): 153–61. Accessed June 10, 2023. https://doi.org/10.1016/S0003-3472(67)80027-7.

Nadelhoffer, Thomas, Jason Shepard, Damien L. Crone, Jim A.C. Everett, Brian D. Earp, and Neil Levy. "Does Encouraging a Belief in Determinism Increase Cheating? Reconsidering the Value of Believing in Free Will." *Cognition*, vol. 203 (2020): 104342. Accessed June 16, 2023. https://doi.org/10.1016/j.cognition.2020.104342.

Nemati, Farshad. "Anthropomorphism in the Context of Scientific Discovery: Implications for Comparative Cognition." *Foundations of Science*, vol. 28, no. 3 (2023): 927–45. Accessed March 30, 2024. https://doi.org/10.1007/s10699-021-09821-1.

Nickles, Thomas. "Bounded Rationality, Scissors, Crowbars, and Pragmatism: Reflections on Herbert Simon." *Mind & Society*, vol. 17, no. 1 (2018): 85–96. Accessed June 11, 2023. https://doi.org/10.1007/s11299-019-00206-3.

Nisbett, Richard E., and Timothy D. Wilson. "Telling More than We Can Know: Verbal Reports on Mental Processes." *Psychological Review*, vol. 84 (1977): 231–59. Accessed June 11, 2023. https://doi.org/10.1037/0033-295X.84.3.231.

Nordli, Samuel, and Peter M. Todd. "Embodied and Embedded Ecological Rationality: A Common Vertabrete Mechanism for Action Selection Underlies Cognition Heuristic Decision-Making in Humans." In *Embodied Bounded Rationality*, edited by Shaun Gallagher, Riccardo Viale, and Vittorio Gallese, 133–43. Lausanne: Frontiers Media SA, 2023.

Nussbaum, Martha C. "Compassion: Human and Animal." In *Understanding Moral Sentiments: Darwinian Perspectives?*, edited by Hilary Putnam, Susan Neiman, and Jeffrey P. Schloss, 123–50. London/New York: Routledge, 2017.

Peirce, Charles S. "Illustrations of the Logic of Science II – How to Make Our Ideas Clear." In *Popular Science Monthly*, (1878): 286–302, Accessed June 15, 2023. https://en.wikisource.org/wiki/Popular_Science_Monthly/Volume_12/January_1878/Illustrations_of_the_Logic_of_Science_II.

Peirce, Charles S. „Niektóre konsekwencje braku czterech zdolności (1868)." In *Charles Sanders Peirce o nieskończonej wspólnocie badaczy*. Translated by Agnieszka Hensoldt, 41–75. Opole: Wydawnictwo Uniwersytetu Warszawskiego, 2009.

Petracca, Enrico. "Embodying Bounded Rationality: From Embodied Bounded Rationality to Embodied Rationality." *Frontiers in Psychology*, vol. 12 (2021): 710607. Accessed June 11, 2023. https://doi.org/10.3389/fpsyg.2021.710607.

Phillips, Webb, Jennifer L. Barnes, Neha Mahajan, Mariko Yamaguchi, and Laurie R. Santos. "'Unwilling' versus 'Unable': Capuchin Monkeys' (*Cebus Apella*) Understanding of Human Intentional Action." *Developmental Science*, vol. 12, no. 6 (2009): 938–45. Accessed November 3, 2023. https://doi.org/10.1111/j.1467-7687.2009.00840.x.

Piekarski, Michał. "Motivation, Counterfactual Predictions and Constraints: Normativity of Predictive Mechanisms." *Synthese*, vol. 200, no. 5 (2022): 352. Accessed June 18, 2023. https://doi.org/10.1007/s11229-022-03837-1.

Pinker, Steven. *Tabula rasa: Spory o naturę ludzką*. Translated by Agnieszka Nowak-Łojewska. Warszawa: Gdańskie Wydawnictwo Psychologiczne, 2012.

Preston, Christopher J. "Animality and Morality: Human Reason as an Animal Activity." *Environmental Values*, vol. 11, no. 4 (2002): 427–42. Accessed June 10, 2023. https://www.jstor.org/stable/30301901.

Redish, A. David. "Vicarious Trial and Error." *Nature Reviews Neuroscience*, vol. 17, no. 3 (2016): 147–59. Accessed June 12, 2023. https://doi.org/10.1038/nrn.2015.30.

Redish, A. David, Samantha V. Abram, Paul J. Cunningham, Anneke A. Duin, Romain Durand-de Cuttoli, Rebecca Kazinka, Adrina Kocharian, et al. "Sunk Cost Sensitivity during Change-of-Mind Decisions Is Informed by Both the Spent and Remaining Costs." *Communications Biology*, vol. 5, no. 1 (2022): 1–17. Accessed June 12, 2023. https://doi.org/10.1038/s42003-022-04235-6.

Reinhold, Annika Stefanie, Juan Ignacio Sanguinetti-Scheck, Konstantin Hartmann, and Michael Brecht. "Behavioral and Neural Correlates of Hide-and-Seek in Rats." *Science*, vol. 365, no. 6458 (2019): 1180–83. Accessed June 10, 2023. https://doi.org/10.1126/science.aax4705.

Rescorla, Robert, and Allan Wagner. "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement." In *Classical Conditioning II: Current Research and Theory*, edited by Artur Black and William Prokasy, 64–99. New York: Appleton-Century-Crofts, 1972.

Ricard, Matthieu, and Wolf Singer. *Beyond the Self: Conversations between Buddhism and Neuroscience*. 1st edition. Cambridge, Massachusetts: The MIT Press, 2017.

Rigoni, Davide, Gilles Pourtois, and Marcel Brass. "'Why Should I Care?' Challenging Free Will Attenuates Neural Reaction to Errors." *Social Cognitive and Affective Neuroscience*, vol. 10, no. 2 (2015): 262–68. Accessed June 15, 2023. https://doi.org/10.1093/scan/nsu068.

Rowlands, Mark. "Animals as Moral Subjects." In *The Routledge Handbook of Philosophy of Animal Minds*, edited by Kristin Andrews and Jacob Beck, 469–74. London–New York: Routledge, 2018.

Rowlands, Mark. *Can Animals Be Moral?* New York: Oxford University Press, 2012.

Rowlands, Mark. *Can Animals Be Persons?* New York: Oxford University Press, 2019.

Ruse, Michael. "Common Sense Morality and Its Evolutionary Underpinnings." In *Scientific Challenges to Common Sense Philosophy*, edited by Rik Peels, Jeroen de Ridder, and Rene van Woudenberg, 160–83. London–New York: Routledge, 2020.

Santolin, Chiara, and Jenny R. Saffran. "Constraints on Statistical Learning Across Species." *Trends in Cognitive Sciences*, vol. 22, no. 1 (2018): 52–63. Accessed August 12, 2023. https://doi.org/10.1016/j.tics.2017.10.003.

Sapolsky, Robert M. *Behave: The Biology of Humans at Our Best and Worst*. London: Vintage, 2018.

Sapolsky, Robert M. *Determined: Life Without Free Will*. London: The Bodley Head, 2023.

Sarkissian, Hagop, Amita Chatterjee, Felipe De Brigard, Joshua Knobe, Shaun Nichols, and Smita Sirker. "Is Belief in Free Will a Cultural Universal?" *Mind & Language*, vol. 25, no. 3 (2010): 346–58. Accessed June 16, 2023. https://doi.org/10.1111/j.1468-0017.2010.01393.x.

Schetz, Adriana. „O tak zwanym problemie prostych umysłów." *Diametros*, vol. 30 (2011), 41–60. Accessed June 10, 2023. https://doi.org/10.13153/DIAM.30.2011.455.

Schloss, Jeffrey P. "Darwinian Explanations of Morality: Accounting for the Normal but Not Normative." In *Understanding Moral Sentiments: Darwinian Perspectives?*, edited by Hilary Putnam, Susan Neiman, and Jeffrey P. Schloss, 81–122. London: Routledge, 2014.

Schwitzgebel, Eric, and Joshua Rust. "The Moral Behavior of Ethics Professors: Relationships among Self-Reported Behavior, Expressed Normative Attitude, and Directly Observed Behavior." *Philosophical Psychology*, vol. 27, no. 3 (2014): 293–327. Accessed April 1, 2024. https://doi.org/10.1080/09515089.2012.727135.

Schwitzgebel, Eric. "Do Ethicists Steal More Books?" *Philosophical Psychology*, vol. 22, no. 6 (2009): 711–25. Accessed April 1, 2024. https://doi.org/10.1080/09515080903409952.

Schwitzgebel, Eric. "How Often Do Ethics Professors Call Their Mothers?" *Aeon*. Accessed June 14, 2020. https://aeon.co/essays/how-often-do-ethics-professors-call-their-mothers.

Schwitzgebel, Eric. "The Moral Behavior of Ethicists and the Role of the Philosopher." In *Experimental Ethics: Toward an Empirical Moral Philosophy*, edited by Christoph Luetge, Hannes Rusch, and Matthias Uhl, 59–64. London: Palgrave Macmillan UK, 2014.

Sellars, Wilfrid. "Philosophy and the Scientific Image of Man." In *Science, Perception and Reality*. Atascadero: Ridgeview Publishing Company, 1991.

Seth, Anil. *Being You: A New Science of Consciousness*. London: Penguin Publishing Group, 2021.

Sevillano, Verónica, and Susan T. Fiske. "Warmth and Competence in Animals." *Journal of Applied Social Psychology*, vol. 46, no. 5 (2016): 276–93. Accessed June 3, 2023. https://doi.org/10.1111/jasp.12361.

Shariff, Azim F., and Kathleen D. Vohs. "The World Without Free Will." *Scientific American*, vol. 310, no. 6 (2014): 76–79. Accessed June 14, 2023. https://doi.org/10.1038/scientificamerican0614-76.

Simon, Herbert Alexander. *Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in a Social Setting*. 1st edition. New York: Wiley and Sons, 1957.

Solomon, Robert C. "Existentialism, Emotions, and the Cultural Limits of Rationality." *Philosophy East and West*, vol. 42, no. 4 (1992): 597. Accessed June 10, 2023. https://doi.org/10.2307/1399671.

Steiner, Adam P., and A. David Redish. "Behavioral and Neurophysiological Correlates of Regret in Rat Decision-Making on a Neuroeconomic Task." *Nature Neuroscience*, vol. 17, no. 7 (2014): 995–1002. Accessed June 11, 2023. https://doi.org/10.1038/nn.3740.

Steward, Helen. *A Metaphysics for Freedom*. Oxford: Oxford University Press, 2014.

Stich, Stephen P. "Could Man Be an Irrational Animal? Some Notes on the Epistemology of Rationality." *Synthese*, vol. 64, no. 1 (1985): 115–35. Accessed June 14, 2023. https://www.jstor.org/stable/20116149.

Strawson, Galen. „O niemożliwości całkowitej odpowiedzialności moralnej." Translated by Jacek Jarocki. *Roczniki Filozoficzne*, vol. 65, no. 1 (2017): 109–29. Accessed June 10, 2023. https://doi.org/10.18290/rf.2017.65.1-6.

Tadeusz Guz – Wykład, 2018. Accessed March 10, 2023. https://www.youtube.com/watch?v=1_LqnK2j_Do.

Tomczyk, Barbara. *Podmiotowość rozszerzonych systemów poznawczych*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, 2022.

Urquiza-Haas, Esmeralda G., and Kurt Kotrschal. "The Mind behind Anthropomorphic Thinking: Attribution of Mental States to Other Species." *Animal Behaviour*, vol. 109 (2015): 167–76. Accessed March 30, 2024. https://doi.org/10.1016/j.anbehav.2015.08.011.

Veit, Fiery, Joe Dewhurst, Krzysztof Dolega, Max Jones, Shaun Stanley, Keith Frankish, and Daniel Dennett. "The Rationale of Rationalization." *Behavioral and Brain Sciences*, vol. 43 (2019): e28. Accessed June 13, 2023. https://doi.org/10.1017/S0140525X19001730.

Viale, Riccardo, Shaun Gallagher, and Vittorio Gallese. "Bounded Rationality, Enactive Problem Solving, and the Neuroscience of Social Interaction." *Frontiers in Psychology,* vol. 14 (2023). Accessed June 11, 2023. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1152866.

Waal, Frans B. M. de. "Anthropomorphism and Anthropodenial: Consistency in Our Thinking about Humans and Other Animals." *Philosophical Topics*, vol. 27, no. 1 (1999): 255–80. https://www.jstor.org/stable/43154308.

Waal, Frans B. M. de. *Bystre zwierzę. Czy jesteśmy dość mądrzy aby zrozumieć mądrość zwierząt?* Translated by Łukasz Lamża. Kraków: Copernicus Center Press, 2019.

Wisniewski, David, Robert Deutschländer, and John-Dylan Haynes. "Free Will Beliefs Are Better Predicted by Dualism than Determinism Beliefs across Different Cultures," edited by Jonathan Jong. *PLOS ONE*, vol. 14, no. 9 (2019): e0221617. Accessed June 15, 2023. https://doi.org/10.1371/journal.pone.0221617.

Yong, Jose C., Norman P. Li, and Satoshi Kanazawa. "Not so Much Rational but Rationalizing: Humans Evolved as Coherence-Seeking, Fiction-Making Animals." *American Psychologist*, vol. 76 (2021): 781–93. Accessed June 13, 2023. https://doi.org/10.1037/amp0000674.

Zimmerman, Aaron Z. *Belief: A Pragmatic Picture*. Oxford: Oxford University Press, 2018.

Zuberbühler, Klaus. "Syntax and Compositionality in Animal Communication." *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 375, no. 1789 (2019): 20190062. Accessed September 9, 2024. https://doi.org/10.1098/rstb.2019.0062.

Zbigniew Słuszkiewicz is a graduate of Philosophy from the Pedagogical University of Cracow (2007). He also completed postgraduate studies in Special Pedagogy at the same institution (2012), Autism, Asperger Syndrome, and Overall Developmental Disorders, Diagnosis, and Therapy (with distinction) at the WSB University of Dąbrowa Górnicza (2015), and Psychology of Crisis and Crisis Intervention at the Jagiellonian University of Cracow (2019). Currently a student of Doctoral School, UKEN, Cracow, An alumnus of the Diverse Intelligences Summer Institute 2021, supported by the Templeton Foundation grant and conducted by the University of California, Los Angeles. In 2022, he was awarded a PRELUDIUM scientific grant from Narodowe Centrum Nauki for a research program titled "The Pragmatic Turn in Embodied Cognition and Mark Rowlands' Category of 'the Moral Subject'." He is an ethics teacher and behavioral therapist at the Centrum Autyzmu i Całościowych Zaburzeń Rozwojowych in Cracow. E-mail: zbigniew.sluszkiewicz@doktorant.up.krakow.pl.

This article was inspired by the cooperation with the Diverse Intelligences Summer Institute, whose programs are financed by the Templeton World Charity Foundation (grant 0333 to the University of California).

Zbigniew Słuszkiewicz jest absolwentem filozofii Uniwersytetu Pedagogicznego w Krakowie (2007). Ukończył również z wyróżnieniem studia podyplomowe na kierunku Pedagogika Specjalna (2012): Autyzm, Zespół Aspergera oraz inne całościowe zaburzenia rozwojowe – diagnoza i terapia (2015) w Wyższej Szkole Biznesu w Dąbrowie Górniczej (2015) oraz Psychologia Kryzysu i Interwencji Kryzysowej na Uniwersytecie Jagiellońskim w Krakowie (2019). Obecnie doktorant UKEN w Krakowie. Jest absolwentem Diverse Intelligences Summer Institute 2021, wspieranego przez grant Fundacji Johna Templetona i prowadzo-

nego przez Uniwersytet Kalifornijski w Los Angeles. W 2022 roku otrzymał grant naukowy PRELUDIUM od Narodowego Centrum Nauki na dwuletni program badawczy zatytułowany „Pragmatyczny zwrot w poznaniu ucieleśnionym i kategoria »podmiotu moralnego« Marka Rowlandsa". Na co dzień pracuje jako nauczyciel etyki i terapeuta behawioralny w Centrum Autyzmu i Całościowych Zaburzeń Rozwojowych w Krakowie.

Artykuł jest częściową realizacją grantu PRELUDIUM 20: Pragmatyczny zwrot ucieleśnionego poznania a kategoria „podmiotu motywowanego moralnie" Marka Rowlandsa (2021/41/N/HS1/04208).