Origins of morality. Kant and modern cognitive science

Natalia V. Zaitseva

Lomonosov Moscow State University, Faculty of Philosophy, GSP-1, Leninskie Gory, 119991 Moscow, Russia

Abstract. The origins of morality, the role of emotions in moral behaviour and other related issues nowadays agitate philosophers as well as natural and cognitive scientists. They investigate the seemingly ‘moral’ and ‘reasonable’ behaviour of social animals, which raises the question of the nature of morality in a new perspective, and thus motivates an appeal to Kant’s moral philosophy. It is noteworthy that Kant did not connect morality exclusively with human nature. Linking moral principles with pure reason, he considered morality as an a priori condition of the moral behaviour of every rational being. In this article, I attempt to compare Kant’s moral rationalism to the ethical doctrine of enactivism. I would suggest that, despite obvious differences, these concepts are quite compatible as indicated by the empirical data of modern life sciences. The result of the combination of these conceptions can be a hybrid position based on the recognition of the necessary character of morality, due to the reasonable structure of the world. In the first section, I consider the Kantian rational interpretation of morality. The second section provides an outline of the enactivist approach. The third part focuses on the contribution that modern natural science can make to the analysis of moral behaviour. In conclusion, I recap major findings and reiterate my position.

Keywords: morality, Kantian ethics, enactivism, evolutionary biology

1 Introduction

We often talk about the social conditionality of human behaviour and readily agree that we are social beings, while still being reluctant to recognise our animality because of the widespread negative connotations of the term. It is widely assumed that being an animal is shameful. Being an animal means being unintelligent and unable to act in accordance with moral principles; also, that only a human being is rational and moral, is a stance with deep cultural, historical and epistemological roots. Kant’s undisputed authority and his practical philosophy seem to have contributed much to spread this opinion. Both in Hegel’s time and today Kant’s moral theory is criticised for being unnatural. It is labelled a “top-down” theory in stark contrast to modern biology-based approaches (cf. Colombetti and Torrance (2009))

* Corresponding author: natvalen@list.ru
ORCID: 0000-0002-1691-1402

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
for the inter-enactive conception of participatory sense-making). In this paper, my aim is not to focus solely on differences between Kantian ethics and modern moral research within the frame of enactivism and evolutionary biology, for example, but to indicate the possibility of a compromise hybrid position between the Kantian *a priori* “top-down” approach and modern “bottom-up” interpretations of morality associated with the biological *a priori* and its natural origin.

Nowadays, there is no doubt about an effective correlation between ideal meaningful objects, such as beliefs, values, emotions etc., which form the content of human consciousness, and their bodily, neural counterparts; accepted, too, is the treatment of the agent as a psychophysical being. No one challenges the conclusion that the brain is a physical correlate of consciousness and cognition cannot be explored without studying the brain and the entire human body as a whole. Modern science has transformed Kant’s concept of experience into the understanding of the world as an integrated cognitive bipolar space, constituted in the process of constant interaction of the agent with environment. Kant was one of those philosophers who brought us closer to understanding a necessary relationship between the thinking person and the sensible world, permeated with rationality: “Thoughts without content are empty, intuitions without concepts are blind. It is thus just as necessary to make the mind’s concepts sensible (i.e., to add an object to them in intuition) as to make our intuitions understandable (i.e., to bring them under concepts). […] The understanding is not capable of intuiting anything, and the senses are not capable to thinking anything. Only from their unification can cognition arise” (*KrV*, A 50-51 / B 74-76; Kant, 1998, pp. 193-194).

In support of my point of view, I begin by considering Kantian moral rationalism and subsequently taking on the enactivist interpretation of morality. Finally, I consider evolitional biology with respect to morality.

## 2 Morality from a rational perspective

While cognitive science generally shares Kant’s idea of the correlativity between the ideal (identical) and sensible (temporal) moments in cognition, cognitive scientists still discuss Kant’s ethics. Perhaps the only thing that both Kant’s followers and proponents of the natural origin of morality agree on is deeming morality a necessary condition for the social agents’ behaviour, as something that lies at the very foundation of social experience and is a condition for moral behaviour. In Kant’s ethics, morality is an *a priori* law whose absolute necessity obliges me to act regardless of my inclinations, constantly evolving feelings and desires. Such changing feelings and emotions cannot oblige us to act morally, so there is a gap between inclination and moral action. Kant contrasted morality and emotions; in his theory, moral action always is an action contrary to inclination – it is not performed in accordance with the moral law, but rather for its sake. Feelings and emotions cannot form a sufficient condition for moral actions. There is also a need for the awareness of duty and a decision to follow it. For Kant all moral concepts are *a priori*; they are absolutely necessary in contrast to sensory natural experiences.

Morality as a formal *a priori* determines the possibility of moral behaviour for all intelligent beings. Kant characterises the moral law, puts emphasis on its absolute necessity as it is in itself, as contrasted with the necessity for something. This is how Kant’s antipsychologism manifests itself. Reason does not perform exclusively in the human world; its legislative and regulatory function can also be exercised outside of it. In addition, unified reason as the lawgiver of the world determines the concept of the human world as a whole, and not as a conglomerate of its various regions, of sensible and intelligible.

Morality as identical and necessary is opposed to sensuality as changeable and accidental. Kant assimilates it to the human capacity for awareness and free will. At the same time, one
should not think that a person ceases to be considered as a whole. A human being is an indissoluble unity of the sensible, external and spiritual internal, so it is impossible to argue that one part of the human has free will, while the other has not, or that one part feels and the other reasons. It was important for Kant to emphasise that moral choice is always a conscious choice, and the decision to follow duty is not based on natural inclinations and desires which are ultimately determined by biological expediency. Natural inclination can serve only as an initial impulse to the exercise of an act, but not as its basis. Awareness as ‘being conscious of’ characterises a moral act as being motivated from within, independently of external natural determinations. No one can force a person to perform moral actions and moral behaviour cannot be imposed. It is the individual himself that makes the decision to act in accordance with the moral law, ‘following the path of duty’. This is the way the individual’s freedom manifests itself. Kant’s juxtaposition of morality and propensity inevitably raises the question as to how ideal identical timeless concepts, a priori entities, correspond to the individual consciousness of a particular temporary corporeal person.

How can we move from sensibility to reflection without breaking a person into two autonomous parts? And what Kantian concept can serve as a mediator for us? Such an emotion-laden concept for Kant is a sense of duty; it concerns a specific feeling that appears in a reflection. The point at issue is a specific feeling that accompanies reasoning, a certain empirical subject’s experience of an ideal entity. The “sense of duty” already reports on a transformation of the ideal, absolute idea of duty, comprehended by reason, into something subjectively given. In the case of a moral choice, the emotional component of the sense of duty manifests itself in satisfaction, in the pleasure stemming from the performed act.

In the Critique of Practical Reason, Kant characterises the sense of duty in terms of the substantives Gefühl or Sinn (sense or feeling). In so doing, he addresses a subjective performance of duty that evokes a “sense of self-satisfaction” (ein Gefühl der Zufriedenheit mit sich selbst). However, it would be wrong to interpret Kant’s ideas through the lens of psychologism. Sense of duty is a moral sense from which the concept of duty is not derived. Free will is directly determined by a moral law, and the concept of duty is only conceivable by reason. Otherwise, Kant argues, “we should have to think of a feeling of a law as such and make what can only be thought by reason an object of sensation” (KpV, AA 05: 39; Kant, 2015, p. 35).

Kant realised that a human being cannot stop being an emotional, natural one, even if she acts rationally. Modern science has experimentally proved the involvement of emotions in cognitive processes. Thus, for example, processing in the prefrontal cortex, part of the brain in charge of cognitive control, is determined by the level of the neuromodulator dopamine (Ott, Jacob and Nieder 2014). It’s nice to do good, and it’s also nice to know that you did the right thing. According to the embodied approach, abstract and concrete concepts are processed by different types of simulations (cf. Borghi and Barsalou, 2019; Buccino, et al., 2016, to mention but a few). Specifically, affective states are essentially involved in simulation representations and processing of abstract concepts and corresponding linguistic expressions (Ponari, Norbury and Vigliocco, 2020) and, what is more, there are some empirical data in favour of the coincidence of brain areas that support both language comprehension and emotions (Pulvermüller and Grisoni, 2020; Mannert, et al., 2019).

At the same time, ideal conditions of experience, including the moral a priori in the Kantian sense, cannot be investigated empirically. There is an inevitable transformation of concepts in the scientifically-based view of morality. Thus, in modern biology, one prefers to talk about built-in, embedded and embodied mechanisms as in the case of neural dispositions in the brain (Damasio, 1999), which are activated as a reaction to a certain stimulus. These mechanisms are based on views of the natural world and animals that differ from Kant’s ideas. Biology has made great strides in the last decade. Today, the ability of some apes to form concepts, make judgments, and even learn and use ASL (the American
Sign Language) has been experimentally proven. Many researchers suggest the difference between humans and other animals is quantitative rather than qualitative (Read, 2020; Hummer, 1985). Scientists and experimental philosophers increasingly talk about “proto-thinking”, as well as animal “proto-morals” as embedded and originally given features. Modern science, in Kantian manner, discusses conditions of experience, the only difference being that the focus of scientists’ attention is on agents of animal nature. Even a cell is considered a separate organism, purposefully acting in its surrounding environment. The purposefulness of organisms’ behaviour is associated with the orientation to a pattern and correction of their actions in cases of violation of the initially set norm. In the case of social animals, proto-morality is associated with ‘ideas about the norm’ as a natural regulator of relationships both within the species and outside it, in inter-specific communication.

As for the Kantian vision of the moral act, it rests upon the understanding of the duality of human nature, which includes a layer of animality. A moral action cannot be absolutely moral, so Kant speaks of a certain degree of morality in relation to human behaviour. A moral act is always motivated by following the idea of duty, but it can never be a pure realisation of it. Our sensibility does not allow us to be absolutely moral. It plays a specific role in our behaviour, being the potency of moral action, but not the moral act itself. It is awareness that makes an act moral.

From the perspective of the scientific data of his day, Kant deemed that an animal is not capable of conscious action nor is it capable of reflection. Awareness, the ability to reflect, characterises intelligent beings. It is important that Kant is not referring to human morality or human reason; he considers moral creatures in general, morality in general, reason in general, i.e. pure reason, realising its intentions in the world of intelligent beings.

The Kantian metaphysics of morality is antipsychological, which distinguishes it from psychological moral conceptions that consider solely intra-specific relations, i.e. the relationships between human agents. At the same time, it is based on a fundamentally different understanding of the world from modern enactivism, for example, not as initially given and significant for the agent, but as a potency for the implementation of the ideas of the reason.

Based on the phenomenological concept of cognition, enactivists start from an understanding of the world as given and strive to avoid unfounded scientific assumptions. They recognise the correlation between ideal-real moments in cognition and do not focus on the difference between intelligent beings and animal agents, but on the search for commonalities between them. But these commonalities are connected primarily with corporeality and environment. Biology clearly demonstrates how changes in the environment alter conditions and prescriptions of behaviour. For Kant, human agents constitute themselves on rational principles, regardless of both natural inclinations and the world in general. According to the enactive paradigm, the living world is initially significant for the agent and, accordingly, initially value-loaded.

3 An enactive account of morality

Nowadays, biology claims to be the metaphysics of nature. It has reached the level that allows raising questions and offering solutions concerning the foundations of the moral and other values that would traditionally belong to the scope of philosophy. Antonio Damasio, Humberto Maturana and Francisco Varella address the question of the biological basis of behaviour that inevitably includes ethics in the metaphysics of nature. Interestingly, Edmund Husserl, the founder of phenomenology and an ideological follower of Kantian transcendentalism, foresaw this development in his unpublished works. He puts forward an idea of a biological generative *a priori* as a starting point for the human being: “here we have
the a-priori of the body’s instincts, originary drives (\textit{Urtriebe}), which bring to fruition (eating, mating etc.) the a-priori itself” (Husserl, 2013, p. 8).

Championing the idea of the world as a unity realised in the interaction between the organism and the environment, modern enactivists emphasise that it is the environment that provides us with an opportunity for moral action. Morality is realised in the sensibly perceived world by underlying it. Sensibility, the flesh, are necessary conditions for an embodied action. Like Kant enactivists believe that we perceive the world as conscious and corporeal beings. But while Kant takes moral freedom and responsibility beyond the limits of perceptual knowledge, enactivists view perception as an integral part of moral life and moral freedom of action.

The distinction between enactivism and Kant’s moral philosophy lies in the difference in the value assumptions underlying the theories. Kant’s top-down theory of morality is based on the value of reason and the autonomy of rational moral knowledge, its independence from natural inclination due to biological expediency. Enactivism is primarily a theory of life in which sensory perception, action and cognition are considered to be inextricably linked. According to enactivism, it is in this relationship that the source and foundation of life, the moral and other values find themselves. Autopoietic enactivism is based on the understanding of embodied reason and the consideration of living organisms as natural self-governing entities. It is a new naturalistic approach to embodied cognition. From an enactivist’s point of view, the philosophy of mind should include ethics and be based on the understanding of an inseparable unity of reason and the organism. When an organism connects with nature, ethics becomes a philosophy of nature.

The founding fathers of enactivism (Varela, Thompson and Rosch, 1991) challenge the value-neutral naturalism of Kantian philosophy. From their perspective, even the simplest metabolic processes of nutrition and independent movement can already be considered in the context of value-based cognition. Living organisms appear to be creative, sense-making creatures. Even ordinary perception is considered as a selective process based on what is valuable to the body. Autonomous living organisms are constitutively dependent on the environment, hence the environment is a value for the organism.

Autopoietic enactivism rests on the rejection of the separation of natural science from the field of values. The theory of embodied knowledge is included in life science. The world is what the animal perceives. Clinging to life, to ourselves, we declare that we value ourselves. The value of life is thus found to lie at the very foundation of the living world. In contrast to Kant, enactivists view propensities for moral values and demands as coming directly from the world itself. The natural world appears in this interpretation to be initially value- and moral-loaded. In animals, constitutive principles and norms are associated with an instinctive community. The animal’s instinct contains its will and the law of causation. Instincts determine what an animal does and why it does it. According to Kant, the sensibly perceived world is value-neutral. It is we who endow the world with values that are the results of practical rational processing. Enactivism, which considers living organisms as self-governing systems conditioned by a perceptual-cognitive relationship with the environment, is based on epistemological premises different from the Kantian approach to the world and morality. The basic moral value is life and its preservation.

The position of enactivism is confirmed in modern biology and neuroscience. Thus, analysing the self-regulating behaviour of the cell, Damasio (2010, p. 36), the renowned biologist, notes that its nucleus and cytoplasm that make up the cell perform complex calculations together, aimed at keeping the cell alive in constantly changing conditions: “obviously, the cell components carrying out those adaptive adjustments were put into place and instructed by the cell’s genetic material”. He is committed to proving the founding role of unconscious management of life processes, making assumptions about the existence of a scheme (disposition), according to which attitudes and intentions of a conscious psyche were
subsequently built. The dispositions are “know-how formulas that code for something like this: if hit from one side, move in the opposite direction for X number of seconds, regardless of the object hitting you or of where you are” (ibid., p. 105), which immediately evokes associations with the Kantian concept of “Schema”. Our conscious psyche brings this basic know-how of life processes management into the realm of knowledge. In fact, in this manner he refers to the rules, patterns and norms that determine the living organism’s development and functioning. Damasio connects dispositions with the concept of necessity, claiming that they do not depend on a person’s will and desires, which implies an emotional program independent of our will.

The enactivist moral theory is sometimes labelled as a naturalistic alternative to Kantian ethics. Guided by a scientific understanding of the sensory-motor nature of perception, enactivists treat moral perception as a specific type of perception. By contrast, it is common knowledge that Kant did not examine the role of perception in the moral life.

4 Evolutionary biology and moral behaviour

Kantian ethics contrasts the rational human being to other animals, thereby resting upon the concept of a rational autonomous will. Enactivism stems from an interpretation of a universal animal autonomy. The animal is autonomous as long as it acts as a self-organising being driven by instincts. An animal’s bodily organism carries out common, built-in, instinctive dispositions. Human agents have the same physical and bodily autonomy. Kant contrasted the autonomous will of human beings with the instinctive animal will. This autonomous will manifests itself in the ability to say No to natural inclinations, desires and instincts. The proposed interpretation begs the question whether animals can ‘say No’ and act against their wishes. Kant would have probably answered No to that. Frans de Waal, a contemporary acclaimed biologist, ethologist and evolutionist (cf. e.g. De Waal, 2002; 2016), takes a different perspective on the problem as he tries to argue in favour of a natural evolutionary origin of human morality that deems social animals’ behaviour as proto-moral.

For example, he discusses a direct perception of injustice in animals, compassion based on a bodily empathic perception of the Other and accompanied by a motor reaction, i.e. an action directed at the Other. Modern science often treats empathy as a pre-understanding. According to F. de Waal, empathy turns out to be another form of perception, i.e. a perception of the Other. Considering a spontaneous emotional reaction directed at the Other as biologically rooted, he emphasises that we act in the same fashion we would in relation to ourselves. Compassion is based on an empathic transfer and, since we are dealing with the perceptual-motor complex, perception is always associated with an action. Empathy helps us to perceive someone else’s misfortune and not to hesitate to come to his/her rescue. The intention to help others has deep biological roots. At the end of the twentieth century, a team of researchers from the University of Parma reported a discovery of the so-called mirror neurons, which are now often considered a neurophysiological basis of understanding. When we perceive others performing different actions, the brain cells that represent motor activity similar to what we observe or expect from others activate. In this sense, the Other is perceived as a projection of myself and I help others, just as I would help myself.

De Waal emphasises the basic role of emotional irrational perception. He criticises Kant’s idea that following a rule or pattern is solely an ability of our species as intelligent beings and points out that other animals also follow patterns. At the same time, de Waal lacks what Kant deemed a critical requirement, namely the awareness of following the ideal, which allows us to talk about independent, free choice. De Waal argues in favour of an ideal built-in scheme in the animal’s head, according to which it chooses or corrects its line of behaviour. From his perspective, any correction is normative by definition. He emphasises that normativity expresses how animals sense a proper order of things and refers to a specific
embedded and mandatory social behavioural instinct. With regard to self-control and counteracting desires and propensities, his experiments with primates convincingly demonstrate that apes are able to control their emotions by following the norm they have an idea of.

While drawing a conclusion about animals’ inherent inclination to justice, candour and mutual trust, de Waal takes as premises not only his own research of primate behaviour, but also refers to the works of Marc Bekoff, a US zoologist (cf. Bekoff and Pierce, 2009), who analysed games of dogs, wolves and coyotes. All those manifestations of the built-in universal moral idea have nothing to do with logic and rational reasoning. This ‘proto-moral’ irrationality thus transcends the boundaries of a particular biological species.

De Waal is certainly aware of the difference between an intuitive proto-moral behaviour of primates and other animals and a moral rational behaviour of humans, which can lead to the same results. He actually considers the animal moral behaviour potency, proto-morality, that has evolved into human conscious moral behaviour. Both behaviours have common roots but pursue different goals. De Waal views man as a citizen of the world and believes that we need to think about universal moral principles. This brings him to the Kantian level of understanding of morality as a form of rational behavior. He reckons that natural ethics only explains how we arrived at today’s moral state.

The evolutionary approach to morality is criticised as reductionism, primitivisation of the human being, who is allegedly presented as a slave to instincts aimed at the species’ survival. On the other hand, de Waal rightly observes that a person is driven by innate values and emotions that guide his/her behaviour rather than dictate it. Like Kant, he believes that they push us in a certain direction but leave us free to act. Echoing antipsychologism, he notes that if we rested our deliberations on an understanding of morality as a genetically determined intra-specific expediency, we probably would not be able to explain the phenomenon of inter-specific cooperation, help and compassion. Notably, there are many examples of this type of behaviour. He is far from identifying ape behaviour with human moral behaviour, but has found commonalities in the inter-specific behaviour and draws a conclusion about its common, unified natural source.

5 Conclusion

Modern science is changing our understanding of animals and the epistemological status of morality. Some animals’ amazing abilities (to perform intellectual operations, to act against their will, to feel ashamed, to show self-identification etc.), the presence of which is confirmed empirically, takes our understanding of reason, morality and its origins to a new level. The common interpretation of the world as having been reasonably arranged and of moral principles as instituted norms, initially given patterns disclosed in contexts of Kantian ethics and a modern scientific approach, testifies to the fundamental nature and extreme relevance of the concepts in question.

In my previous papers, I developed the idea of a universal nature of cognition-typification as adaptation. I considered cognition as a multi-layered and multi-act cognitive structure whose intentionality determines the interaction between an autonomous living agent and the surrounding world. In this context, the basic emotions of pleasure and displeasure were considered as proto-categories that allow typifying objects in the process of cognition as constituting the world for the agent, i.e. the world as an integral bipolar cognitive space.

A reference to morality invokes a consideration of a world that includes other active agents, i.e. of a common lifeworld. In this case, Kant’s rational theory of morality, De Waal’s conception of irrational proto-morality and the enactivists’ moral account, based on an understanding of an initially value-laden world, are attempts to address the necessary, initially given condition for a common world of social interaction. This condition turns out
to be morality. It is granted the status of law, whose necessity manifests itself in agents’
actions in pursuit of a common good. The tendency to interpret cognition and morality in
terms of community, unity and identity, in my opinion, bridges the gap between Kant’s
rational approach and biologically oriented modern approaches. They appear common in
their antipsychological and normative interpretation of morality.

Comparing top-down and bottom-up theories, we find their essential similarity against
the background of obvious differences in the accepted assumptions. Here and there a
meticulous researcher finds normativity, the necessary character of morality, due to the
reasonable structure of the world. The only difference in this regard is that Kant sees the
source of this normativity in absolute reason, while enactivists and evolutionary biologists
expose it in the structure of a lifeworld.

**Acknowledgments.** I would like to thank the organisers of the 12th Kant-Readings International
Conference “Kant and the Ethics of Enlightenment: Historical Roots and Contemporary Relevance” for
the kind opportunity provided to express my views and position in this follow-up publication. This
research has been supported by the Interdisciplinary Scientific and Educational School of Moscow
University “Brain, Cognitive Systems, Artificial Intelligence”.

**References**


De Waal, F., 2016. *Are We Smart Enough to Know How Smart Animals Are?* New York: *W.W. Norton & Company*.


