What Are the Criteria of Personality in Artificial Intelligence in Relation to Moral Status? *

Saeid Karimi✉ 1 ⓘD | **Morteza Tabatabaie² ** ⓘD | **Mohsen Javadi³** ⓘD | **Saeedeh Momtazi⁴** ⓘD

1. Corresponding Author, PhD Candidate of Moral Philosophy, University of Religions and Denomenations, Qom, Iran. E-mail: saeedsaeid90508@gmail.com

2. Assistant Professor, Department of Moral Philosophy, University of Religions and Denomenations, Qom, Iran. E-mail: morteza.taba64@gmail.com

3. Professor, Department of Ethics, University of Qom, Qom, Iran. E-mail: javadi.m@qom.ac.ir

4. Associate Professor, Department of Artificial Intelligence and Robotics, Amirkabir University of Technology, Tehran, Iran. E-mail: momtazi@aut.ac.ir

Article Info	ABSTRACT
Article type: Research Article **Article history:** Received 18 October 2025 Received in revised form 25 October 2025 Accepted 27 November 2025 Published online 14 February 2026 **Keywords**: Personhood, Moral Status, Artificial Intelligence, Kantian Deontology, Virtue Ethics, Consciousness.	The present article examines the possibility of attributing personhood to artificial intelligence agents, a concept central to determining moral status. The debate on personhood has long been central in applied ethics, particularly in discussions on abortion, where philosophers such as Mary Anne Warren proposed five criteria—consciousness, reasoning, self-motivated activity, capacity for communication, and self-awareness—as key indicators of personhood. The present study applies these criteria to artificial intelligence systems and asks whether their cognitive and functional capacities are sufficient for moral consideration. While certain features such as memory, goal-directed behavior, and limited moral interaction are identifiable in some AI systems, the absence of self-awareness and subjective experience remains a fundamental obstacle to full personhood. The article further engages with the views of Kant, Locke, DeGrazia, and Searle, assessing the possibility of AI's moral standing—whether direct or indirect—through ethical frameworks such as deontology and virtue ethics. It concludes that although attributing personhood to AI remains highly problematic in its current state, addressing this issue is an urgent necessity for contemporary moral philosophy.

Intruduction

In recent decades, the rapid growth of technologies based on artificial intelligence has opened up new horizons in our understanding of agency, decision-making, and interaction. Intelligent agents, now designed as advanced systems, not only possess the ability to process information, learn and adapt to their environment, but in many cases are also capable of linguistic communication, analysing complex situations, and even producing seemingly rational responses. In light of such developments, one of the most fundamental ethical and philosophical questions of our time arises: Can artificial intelligent agents possess moral status? And if the answer is affirmative, what degree of moral responsibility or moral rights would such a status entail?

This question, in turn, leads us to the concept of personhood since one of the most significant criteria often cited for an entity's possession of moral status is precisely the set of characteristics related to personhood. Thus, raising questions such as "Can artificial intelligence be a person?" or "At what point should a non-human agent be considered to have moral rights?" becomes crucial not only from a cognitive standpoint, but also from ethical, legal, and social perspectives. These questions, beyond determining the place of AI agents in our moral framework, also have a direct impact on policymaking, technology-related law, and human behavioural patterns in encountering with such agents.

The concept of personhood is among the most fundamental and, at the same time, most contentious notions in moral philosophy. If an entity is deemed to possess moral personhood, it would inevitably be entitled to specific moral rights that others are obliged to respect (Warren, 1997) It must be noted, however, that the notion of personhood is itself ambiguous. One of the most important fields where the concept of personhood plays a foundational role is applied ethics, particularly the debate over abortion. The central question here is whether a fetus can be considered a person. If personhood is established for the fetus, then abortion would be morally impermissible. Consequently, much of the historical discussion of personhood has been developed in connection with debates on abortion.

At present, the existence of phenomenal consciousness—or what David Chalmers (2022) famously calls the "hard problem" of consciousness—cannot be established in intelligent artificial agents. What may plausibly be ascribed to artificial intelligence is a form of access or psychological consciousness, which is third-personally accessible and open to empirical evaluation. By contrast, first-person consciousness, which pertains to inner subjective experience and direct awareness of one's own mental states, is with certainty attributable only to human beings. On this basis, the present article defends a view of personhood as graded and scalar rather than binary. In other words, intelligent agents may possess a degree of personhood, but not at the level of human beings. In the absence of phenomenal consciousness, their personhood must be considered diminished and, at best, reducible to a form of access-consciousness.

The rapid progress of AI technologies has compelled a reconsideration of traditional concepts associated with personhood, creating the impression that the boundary between humans and robots is becoming increasingly blurred. Some scholars have pointed to advanced AI systems such as Sophia—the humanoid robot that was granted citizenship and even became affiliated with the United Nations Development Programme—as raising the question of whether such technologies might, in the future, attain a form of consciousness (Burckle and Peters, 2024).

To clarify this issue, it is essential to distinguish between two key concepts in moral philosophy: moral status and moral agency. Moral status refers to whether an entity can be the subject of our moral concern—that is, whether we owe it moral duties or obligations. Moral agency, by contrast, refers to the capacity of an entity to understand and perform moral actions, such as discerning right from wrong and acting on moral reasons. This distinction becomes especially relevant in the context of artificial intelligence: even if an AI system cannot qualify as a moral agent, it might still warrant moral status. As Coeckelbergh observes, this leads us to the question known in moral philosophy as moral considerability: can an artificial agent, without itself being a moral agent, nonetheless be the object of our moral duties? (Coeckelbergh, 2020, 48-50).

In this context, the concept of personhood plays a pivotal role. Personhood is not merely a metaphysical debate about identity and self-consciousness; it also carries profound ethical and legal implications. The concept has been extensively examined in applied ethics, particularly in areas such as animal ethics and the moral status of the fetus in abortion debates, where it has functioned as one of the central criteria for determining moral status. These discussions provide a theoretical framework that can serve as a preliminary guide for analysing the moral standing of AI. Thus, a comparative study of the theories of personhood developed in these domains, and their applicability to artificial agents, constitutes a necessary methodological step in the analysis of this novel issue.

According to several perspectives, features such as self-awareness, rationality, the possession of interests, the capacity to experience pleasure and pain, or the ability to pursue goal-directed behavior may serve as criteria that confer moral status upon an entity. Among the most prominent of these features is personhood itself, which has been analysed by philosophers such as John Locke, Immanuel Kant, and, in contemporary times, thinkers like Mary Anne Warren, Tom Regan, and Peter Singer.

The present article is part of a broader research project on the moral status of artificial intelligence. However, in this paper, the focus will be on the concept of personhood as one of the fundamental prerequisites for possessing moral status. In this regard, an effort will be made to examine classical and contemporary theories of personhood and to analyze the possibility of their applicability to the characteristics of artificial intelligent agents, in order to assess whether this concept can be attributed to such entities. It will also be shown that, although personhood can be considered a key pillar in the theory of moral status, there are conceptual, empirical, and normative challenges in extending it to artificial intelligence that require careful analysis.

1. Moral Status

Many debates in applied ethics concerning how we should treat human beings, nonhuman animals, and even the environment, are grounded in theoretical discussions about moral status and its foundation. Although it is generally accepted that all cognitively normal adult humans possess full moral status, history demonstrates that certain social groups—such as foreigners, minorities, women, and persons with physical disabilities—have been deprived of this status or had their moral standing diminished. Moreover, it must be noted that assessing the moral status of certain groups, particularly humans in special conditions such as those with cognitive impairments or at the early stages of biological development, presents additional theoretical complexities (Jaworska and Tannenbaum, 2021).

The term moral status is employed in philosophical literature in at least five distinct but interconnected senses. In its first sense, the term merely refers to the salient moral properties of a given thing; in other words, when the rightness, wrongness, or permissibility of an act or policy is under discussion, one might say that the moral status of that act or policy is at issue. For instance, in titles such as The Moral Status of Abortion, it is this first sense of moral status that is invoked (DiSilvestro, 2010, 17-18).

In the second sense, the concept of moral status represents one of the most fundamental and pervasive notions in ethics. Many contemporary moral controversies, including debates surrounding stem cell research or therapeutic cloning, essentially revolve around the question of whether embryos possess some form of intrinsic importance that entitles them to particular moral consideration. Likewise, debates on abortion and animal rights examine the moral status of fetuses or nonhuman beings. The idea of the moral status of persons plays a central role in many theories concerning basic human rights and duties, and thus lies at the heart of ethical reflection. The key distinction between this and the first sense is that it applies only to concrete entities (and not to actions or policies). Nevertheless, merely claiming that something has moral status—whether in the first or second sense—provides no precise information about the degree or quality of that status, just as saying that something has height is insufficient without specifying its measure (DiSilvestro, 2010, 11).

In the third sense, which is associated with Mary Anne Warren's interpretation of moral status, the term refers to the idea that a being is morally considerable in such a way that moral agents have direct moral obligations toward it. According to Warren, if a being possesses moral status, one cannot treat it arbitrarily; rather, its needs and interests must be taken into account in moral decision-making. This is not merely because doing so might benefit others, but because the needs of that being themselves have intrinsic moral significance (Warren, 1997, 3). In this sense, the claim that a being has moral status points to the existence of certain features in that being—for instance, that it can be the direct object of moral obligations (DiSilvestro, 2010, 18-19).

In the fourth sense, Elizabeth Harman extends the concept of moral status by linking it to the moral significance of being harmed. On her account, a being has moral status if harming it is morally significant in itself. For example, if harm is done to a person like Alice,

that harm constitutes a moral reason against performing the act, solely in virtue of the harm suffered by Alice herself. By contrast, if Alice's car is damaged, the harm is morally relevant only insofar as it bears upon Alice, not because the car itself is an object of moral concern (Harman, 2003, 174). As she formulates it in her dissertation: "If something is ever harmed, then it has moral status provided that we have reasons against harming it, simply because of the badness of the harm for it" (DiSilvestro, 2010, 15).

DiSilvestro subsequently introduces the notion of serious moral status as a construct that integrates elements of both Warren's and Harman's approaches, while extending their scope and moral implications. Possessing such status entails not only that harming or mistreating the being is morally prohibited, but also that even slandering or cursing it carries strong moral constraints. Such a being is not only deserving of respect but of justice, and there is always a moral reason in favour of benefiting it. Serious moral status, therefore, refers to a set of features typically possessed by ordinary adult human beings—for instance, that there is a strong moral presumption against killing them (DiSilvestro, 2010, 11).

The contrast among the interpretations of Warren, Harman, and DiSilvestro highlights an ongoing effort to provide a coherent foundation for moral evaluation of beings. Warren, by stressing "moral considerability" and "direct moral obligations" toward beings, conceives moral status as dependent on certain intrinsic or functional properties of those beings that generate obligations even in the absence of benefits for others. Harman, in contrast, grounds her analysis in the moral significance of harm itself: the fact that harm to a being, simply qua harm, gives rise to a reason to avoid it. While this perspective may be compelling within empiricist frameworks, it requires further elaboration to clearly delineate between beings that possess intrinsic moral value and objects whose significance arises only through their relation to human agents.

In introducing the notion of serious moral status, the authors seek to establish a connection between the aforementioned interpretations and a more expansive view—one that emphasizes both the necessity of avoiding harm and the intrinsic rights and respect owed to individuals. What becomes central here is the reference to a "set of typical human capacities" as the foundation of this status; qualities such as self-awareness, rationality, the ability to have interests, or to act purposively. Although this criterion appears to proceed on the basis of a kind of moral realism and anthropocentrism, it nonetheless faces the risk of excluding certain human beings (such as infants, persons with cognitive disabilities, or patients in persistent vegetative states), unless one adopts an expanded conception of "typical human capacities" that also incorporates potentiality. Overall, these analyses demonstrate that the concept of moral status—despite its seemingly straightforward character—is imbued with complex epistemological and metaethical tensions that make the determination of its scope one of the most challenging issues in applied ethics. At this point, a precise and logically consistent definition may be provided in light of DeGrazia's account:

> To say that something, X, has moral status is (1) to say that moral agents
> have obligations regarding how they treat X, (2) that X possesses interests

(or a welfare), and (3) that X's interests provide the basis of the relevant obligations.

An alternative formulation, according to DeGrazia, is that X has moral status if and only if:

(1) moral agents have obligations regarding how they treat X, and (2) they have these obligations for X's sake.

A simpler formulation is to equate moral status with intrinsic moral value—provided that one assumes the bearers of such value also possess interests or a welfare of their own (DeGrazia, 2008, 183).

2 The Concept of Personhood: Features and Criteria

One of the significant criteria for beings to possess moral status is personhood. Indirectly, it can be said that in discussions about the concept of "being a person," a set of biological, psychological, rational, social, and legal characteristics are considered. At the biological level, attributes such as being human, possessing a specific genetic structure, having bodily organs, and basic abilities such as movement, breathing, eating, and sleeping is relevant. Alongside these, psychological dimensions such as consciousness, feeling, perception, and the ability to use language, tools, or symbols are also taken into account. Rational features such as reasoning, abstraction, learning from experience, and future orientation are likewise regarded as crucial criteria. From a social perspective, the capacity for empathy, cooperation, understanding others, and engaging in interaction holds significance. Finally, legal aspects also shape personhood, including recognition under the law, the ability to enter contracts, holding citizenship, and having property rights. Collectively, these factors contribute to our definition of a person, even though no single criterion is sufficient on its own (English, 1975, 235) From Locke's perspective, a person is defined as:

> For since consciousness always accompanies thinking, and 'tis that, that makes every one to be, what he calls self; and thereby distinguishes himself from all other thinking things, in this alone consists personal Identity, i.e. the sameness of a rational Being (Locke, 1975, 335).

Locke's statement here offers one of the most influential definitions of personhood in the history of philosophy. He stresses that a person is not only a thinking being but also capable of recognizing itself as itself. This means that a person not only thinks but also knows that it is thinking, and this self-awareness is made possible through inner experience and direct awareness of its own mental acts. Importantly, for Locke, being a person is not realized merely in the present moment; it requires the mind's ability to consciously connect past, present, and future times. This aspect of Locke's account underscores the role of psychological continuity in defining personhood.

According to Locke, as long as a being cannot recognize itself as the same entity across different times and places, it cannot be considered a person. Therefore, animals lacking this level of self-awareness do not qualify as persons. Moreover, if artificial intelligence were to

develop a psychological structure capable of maintaining self-awareness over time and referring back to its past actions, one might ask whether such intelligence could qualify as possessing personhood. However, for Locke, mere intelligent performance or machine learning is insufficient for personhood; the necessary condition is the sustained self-consciousness of being the subject of experience.

Despite possible similarities between humans and nonhumans in certain capacities, there is a distinctive feature that sets humans apart—one that cannot be reduced merely to observable behaviours indicative of specific capacities. This distinctive feature is expressed in the concept of personhood. It is often argued that personhood delineates a group of beings that matter morally, a category that closely overlaps with being human. In the history of thought, this view is especially associated with Kant, who is among the most prominent philosophers to argue that personhood is the fundamental quality that bestows moral worth upon a being and renders it morally considerable (Gruen and Monsó, 2024). Kant writes:

> the human being, and in general every rational being, exists as end in itself, not merely as means to the discretionary use of this or that will, but in all its actions, those directed toward itself as well as those directed toward other rational beings, it must always at the same time be considered as an end. All objects of inclinations have only a conditioned worth; for if the inclinations and the needs grounded on them did not exist, then their object would be without worth. The inclinations themselves, however, as sources of needs, are so little of absolute worth, to be wished for in themselves, that rather to be entirely free of them must be the universal wish of every rational being.60 Thus the worth of all objects to be acquired through our action is always conditioned. The beings whose existence rests not on our will but on nature nevertheless have, if they are beings without reason, only a relative worth as means, and are called things; rational beings, by contrast, are called persons, because their nature already marks them out as ends in themselves, i.e., as something that may not be used merely as means, hence to that extent limits all arbitrary choice61 (and is an object of respect) (Kant, 2002, 45-46).

According to this principle, personhood is precisely rooted in this capacity for rationality and autonomy. In other words, a rational being is someone who can legislate moral laws for itself and accept responsibility for them; and this very feature makes it an object of moral respect. In contrast, non-rational beings (such as animals or objects) have only relative value—that is, a value derived from their usefulness for the desires and goals of others. In light of this view, one may argue that animals, being devoid of autonomy and practical reason, lack moral personhood from Kant's perspective and should not be considered ends in themselves, although they may still be considered indirectly in ethical deliberations.

Based on Kantian principles regarding artificial intelligence, the main criterion for possessing moral personhood is the ability for genuine autonomy, practical reason, and

action according to the categorical imperative. If an AI system could not only make decisions but also determine the basis of its decisions autonomously according to moral principles, then one could ask whether it is deserving of the status of a person. In this way, Kant's text forms not only one of the most prominent defenses of the moral status of humans in the foundations of ethics but is also widely applied today in moral philosophy for analyzing the ethical status of artificial intelligence, animals, and even embryos.

Christine Korsgaard, a contemporary American philosopher and one of the foremost interpreters of Kantian ethics in the modern era, emphasizes the reflective nature of the human mind without claiming full and certain access to its contents. Unlike some philosophers who consider the mind entirely transparent, she, like Kant, holds that we are as uncertain about our own thoughts and feelings as we are about external matters. However, she identifies the distinctive feature of the human mind as its capacity for self-reflection. This capacity allows us not only to be aware of our perceptions and desires but also to question them, distance ourselves from them, and make judgments about them. Consequently, a fundamental issue arises, known as normativity for humans: whether a perception or desire we have genuinely provides a reason for belief or action. The reflective mind cannot merely be swayed by desire or perception; it must have a reason capable of withstanding reflection (Korsgaard, 1996, 92-93).

This conception of the human mind as a reflective being is deeply connected to the notion of personhood. A person is not merely a being capable of perception and desire but one who can step back from these mental states, critique them, and act on reasons. This capacity for reflection and evaluation is precisely what confers a normative status upon a being—that is, a being that is accountable for its actions and can be expected to provide rational justifications or ethical reasoning.

Regarding non-human animals, Korsgaard clearly distinguishes them from persons: animals may possess conscious perceptions and desires, but they lack the reflective mental structure that allows them to question their desires. They may act under the influence of desire, but they cannot ask themselves, "Should I do this?"(Korsgaard, 1996, 93)

In the context of artificial intelligence, the crucial question is whether such systems can achieve a level of self-awareness and reflectivity that allows them to question their own reasons. If they merely follow algorithms and data without the ability to distance themselves from their decision-making processes, then, however complex and efficient they may be, they still do not possess personhood. Personhood requires having a "reflective mind": a mind that not only acts but also deliberates on the rightness or wrongness of its actions. In other words, what transforms a being from a thing into a person is not its computational power but its ability to stand back from impulses and reflect on their value. Consequently, ethical evaluation or ascription of intrinsic dignity to beings must be based on their capacity for normative reflection, not merely on cognitive ability or behavioral performance.

David DeGrazia, an American philosopher and a leading figure in applied ethics, particularly bioethics and animal ethics, argues that the concept of "personhood" as a basis for moral status is only useful if it is defined unambiguously and its relation to moral status

is explained coherently. He notes that many philosophers adopt similar psychological conceptions of personhood, such as John Locke's view, which defines a person as a being with self-awareness over time and with rational and reflective capacities. According to such views, neither fetuses nor infants count as persons. However, DeGrazia maintains that appeals to personhood are often ineffective for justifying moral status because the criteria for personhood are disputed and its connection to moral status remains unclear. He emphasizes sentience as a concept that has a clear link to moral status and is both a necessary and sufficient condition for having interests. Therefore, he asserts that we should only appeal to the concept of personhood when it is precisely and usefully defined and its relationship to moral status is logically explained (DeGrazia, 2021, 44-45).

In this regard, DeGrazia strongly criticizes the use of personhood to justify moral status unless the concept is clearly defined and its connection to moral status is logically and systematically articulated. In other words, he stresses that acceptance of such a concept requires assurance that the definition can serve as a basis for reasonable moral evaluation, rather than merely functioning as an anthropocentric or species-centered tool to defend human rights over other beings.

Mary Anne Warren argues that neither mere life nor sentience alone can successfully serve as the sole criterion for determining moral status. She notes that many philosophers have turned to the concept of personhood in search of such a criterion, as it seems to play a fundamental role in justifying the strong moral status that humans typically ascribe to themselves and their close associates. According to Warren, definitions of personhood can be categorized into two types: maximalist definitions, which regard moral agency or at least the potential for it as necessary for personhood, and minimalist definitions, which consider mere capacity for thought and self-awareness sufficient. Warren explains that Kant's definition of personhood is a maximalist one: rational and moral agency is both necessary and sufficient for full moral status. She calls this the "person-only view" and criticizes it as inadequate, because sentient beings without moral agency can still possess moral status. Accordingly, she proposes the "person-plus view," arguing that while moral agency is a sufficient condition for full moral status, it is not a necessary one. Therefore, some sentient beings that lack moral agency may still be entitled to full moral status (Warren, 1997, 90).

In another section, Warren analyses Tom Regan's view. According to her, Regan presents another version of the "personhood-only view," one that is based on a minimalist definition of personhood. Regan maintains that all beings that are "subjects-of-a-life" — and probably only such beings — possess moral standing, and that this standing is equal among them. For Regan, normal mammals over one year of age qualify as subjects-of-a-life and therefore have moral rights equal to those of humans. Warren notes that this version of the personhood-only view, while granting strong moral standing to many sentient beings, nonetheless denies moral standing to many others, particularly non-sentient organisms, biological species, and non-living elements of nature. Ultimately, Warren argues that being a subject-of-a-life cannot serve as the sole valid criterion for determining moral standing (Warren, 1997, 90-91).

Warren further explains that defining personhood is far more difficult than defining life or sentience, because there is a strong conceptual link between personhood and the possession of full moral standing. She points out that in moral debates, advocates of equal moral standing for some animals often call those animals "persons," while their opponents reject such a claim. Similarly, opponents of abortion often hold that human embryos count as persons from the moment of conception, whereas defenders of women's right to choose argue that embryos attain personhood only at later stages of development (e.g., at viability or birth). Warren stresses that these disagreements are not merely about differences in belief regarding the mental capacities of beings, but also about prior disagreements concerning their moral standing. She also observes that some philosophers have concluded that the term "person" is essentially an honorific, indicating the special moral standing of certain beings without necessarily referring to observable empirical traits such as life, sentience, or rationality. From this perspective, personhood is thus more of a moral evaluation than an empirical description (Warren, 1997, 91).

Furthermore, some philosophers — especially those who maintain that only human beings can be persons — advance the view that personhood is exclusively tied to human beings. In this regard, Saponzis argues that in ordinary language, the term "person" necessarily refers to a being with a human body. He emphasises that even if the behaviour of an adult dog in terms of intelligence and self-awareness surpasses that of a human infant or a severely disabled adult, the dog would still not count as a person, while those humans would be regarded as such. According to him, without a human body, no being can be considered a person (Warren, 1997, 92).

Warren argues that the common concept of "person" is less tightly linked to having a human body than Sapontzis assumes. He points out that in children's books, animals are often presented as persons who speak like humans, wear clothes, and drive cars, indicating that the idea of non-human persons is entirely conceivable. According to him, in religious traditions as well, beings such as gods, goddesses, spirits, and angels—even if they lack a human body or possess an animal-like or hybrid appearance—are considered persons, since traits like individual personality and the ability to engage in intelligent dialogue with humans are attributed to them (Warren, 1997, 92-93).

In a critical article, Sapontzis warns against the increasingly broad application of the concept of "person" in theorizing and ethical decision-making. He contends that the concept is, on the one hand, logically ambiguous and, on the other hand, ethically contestable. He emphasizes that in common usage, two distinct interpretations of "person" are wrongly conflated: one is the moral concept of person, which is connected with ethical valuations such as the assignment of rights, duties, and respect, and the other is the metaphysical concept of person, which is understood as a kind of entity within our experiential structure of the world. Sapontzis stresses that the moral concept pertains to social and legal responsibilities, whereas the metaphysical concept mainly concerns descriptive features of beings in the world. From a metaphysical perspective, a "person" is an entity that possesses a body, life, emotions, the capacity for purposeful action, and the ability to form ideas about

the world. In everyday life, these beings are identified as humans and are distinguished from inanimate objects, machines, plants, animals, and spirits. This distinction is based on a combination of bodily form and coherent, self-initiated purposeful behavior (Sapontzis, 1981, 607 -608).

In his critique of the concept of "person," Sapontzis emphasizes that in everyday experience, identifying individuals as persons is not based solely on behavioral features such as consciousness or rationality; bodily form also plays an essential role. He explains that even if the behavior of an adult dog, in terms of self-awareness or intelligence, surpasses that of a human infant or an individual with severe mental and physical impairments, it would still not be considered a person due to the absence of a human body. In contrast, humans with minimal behavioral capacities remain classified as persons by virtue of possessing a human body. He further notes that in popular culture and in contexts such as discussions of abortion, the human form of the fetus plays an important role in judgments about personhood. According to him, our understanding of "person" in this metaphysical sense is tied to the human body, and conceptually, the term "person" applies only to humans (Sapontzis, 1981, 608).

In this discussion, Warren and Sapontzis represent two different approaches to the concept of "person." Warren emphasizes the possibility of conceiving non-human persons and shows that in culture and religious traditions, beings with non-human appearances but endowed with traits such as self-consciousness and the capacity for dialogue are accepted as persons. In contrast, Sapontzis takes a critical view of the expanding use of the concept of "person," considering it both logically and ethically problematic. He distinguishes between the moral and metaphysical interpretations of the concept and emphasizes that in lived human experience, the human body is a key element in identifying a person, not merely behavioral or mental features. Therefore, the main difference between these two perspectives lies in the degree to which the concept of "person" depends on having a human body: Warren rejects this dependency, whereas Sapontzis considers it essential.

Warren presents two types of definitions of a person: Maximal Definition: According to this view, for an entity to be a "person," it must possess the capacity for moral agency. That is, the entity must be able to make moral decisions and act according to ethical principles. Within this framework, an artificial intelligence system would be recognized as a "person" only if it could act according to moral principles (e.g., in complex ethical situations). Minimal Definition: In this approach, the criterion for personhood goes beyond moral agency and requires only self-awareness, logical reasoning, or mental capacities. From this perspective, if an artificial intelligence system (even without the ability to act morally) is capable of understanding, thinking, and experiencing itself, it may be recognized as a "person" (Warren, 1997, 94-95).

Thus, if artificial intelligence reaches a level of self-awareness and cognitive capacity, even if it cannot fully act morally, it could potentially hold moral rights comparable to humans. In his theory of "enhanced personhood," Warren explains that moral agency may be a sufficient condition for granting moral status to an entity, but it is not a necessary one.

This means that even if an entity (human or non-human) cannot fully act as a moral agent, it may still possess moral rights based on its mental capacities or self-awareness. Within this framework, artificial intelligence could be recognized as a morally significant being if it possesses traits such as self-awareness, the understanding of emotions, or the capacity to experience suffering, even without full moral decision-making capabilities. This view is particularly applicable to AGI (artificial general intelligence) systems that attain a certain level of awareness and cognition.

One important point in the discussion of the moral status of artificial intelligence, similar to what is often raised in science-fiction stories, is that we do not limit personhood solely to human beings. In many fictional stories, non-human entities (such as aliens or even self-aware machines) are initially not recognized as "persons," but through interaction and the understanding of these beings' mental capacities, members of the human community come to realize that these entities also possess personality, emotions, and even self-awareness, and therefore should have rights comparable to those of humans.

This concept also arises in the real world: a self-aware artificial intelligence that, through interaction with humans, comes to understand emotions and experiences, may be considered a "person" by many humans. Even if these entities lack biological material or a human body, they may still morally require protection and rights comparable to humans. If artificial intelligence reaches a level of self-awareness and emotional capacity, according to Warren's view, mental capacities and awareness can serve as criteria for attributing moral status to it. This could imply moral rights for artificial intelligence, even if its ethical capabilities are not fully developed. In other words, artificial intelligence may require moral and legal protection similar to living beings such as humans and animals, particularly if it is capable of experiencing emotions or suffering.

It is important to distinguish between the concept of "moral personhood" and "moral status." On this basis, it is possible for an entity to possess moral status without being considered a moral person. The reverse, however, does not hold: if an entity possesses moral personhood, it necessarily has moral status as well (a relation of universal to absolute particular). In other words, moral personhood constitutes a stronger and more complete form of moral status, requiring certain features such as rationality, self-awareness, or the capacity to understand moral responsibility. Therefore, all entities with moral personhood undoubtedly possess fundamental moral rights and considerations, while other entities may possess a degree of moral status solely because of capacities such as sensitivity to pain or the ability to experience suffering, without being considered moral persons (Gordon, 2020).

As noted, "personhood" is a foundational concept in determining the boundaries of responsibilities, rights, and moral values, and it plays a central role in clarifying which entities are deserving of moral consideration. Since different definitions of this concept have been offered across various philosophical traditions, examining the views of Aristotle and Kant can introduce us to two distinct yet influential approaches in this area.

3. Analysis of Personhood from Kantian Deontological Perspective

Immanuel Kant writes: "Two things fill the mind with ever new and increasing admiration and reverence, the more often and the more steadily we reflect on them: the starry heavens above me and the moral law within me (Kant, 2015, 129).

According to Mary Anne Warren, Kant posits that what makes a human a person is the presence of an "inner moral law." This moral law forms the basis of an individual's moral worth, such that only persons can have moral duties and only toward them can moral obligations be ascribed. Consequently, beings that are not considered persons are merely of instrumental value and are treated as objects. While Kant maintains that we should not treat sentient animals cruelly, this duty pertains not to the animals themselves but to other persons who might be harmed by such behavior (Warren, 1997, 96).

Kant's view of personhood does not precisely distinguish all humans as morally considerable beings. The primary issue is that some humans, such as infants and individuals in comas, lack the rational and self-reflective capacities associated with personhood. This issue, known as the "marginal cases problem," leads to many beings who are morally considerable being excluded from moral consideration under this criterion (Gruen and Monsó, 2024).

Kant believes that humans have no direct and immediate duties toward animals because they lack self-awareness and rationality, and thus cannot be regarded as morally responsible agents. From this perspective, the question of why humans exist is meaningful, but the same question does not apply to animals. However, Kant emphasizes that we have indirect duties toward animals; our treatment of them reflects our behavior toward humans. He argues that because animals are comparable to humans in many natural aspects, our treatment of them indicates our level of humanity and benevolence toward others. For instance, if an animal exhibits loyal and service-oriented behavior, caring for and respecting that animal can be a sign of upholding human dignity. Therefore, if someone harms an animal, even though the animal lacks moral judgment, such behavior can damage the human spirit and make the person cruel toward humans as well. Kant also refers to an educational theme, stating that humans should learn from childhood that kindness to animals is a way to cultivate moral feelings. He cites the example of Hogarth, who shows how cruelty to animals in childhood may lead to crimes and murder in adulthood. Thus, fostering affection and compassion for animals is necessary not because of animal rights but to preserve human dignity and humanity (Kant, 1930, 239-241).

Kant, by emphasizing humans' indirect duties toward animals, focuses not on the direct rights of animals but on the moral education of humans. He argues that human behavior toward animals is considered a reflection of an individual's moral standing in interactions with other humans. In other words, virtuous treatment of animals is not only relevant to the animals themselves but also reflects a form of self-awareness and benevolence that can influence human relationships.

This perspective can similarly be extended to analyses concerning artificial intelligence (AI). In applying Kantian thought to AI, although AI lacks human-like self-awareness and consciousness, human behavior toward AI systems can still reflect an individual's ethical character and social responsibilities. From this perspective, interactions with AI should demonstrate humanity and moral responsibility. Just as with animals, human conduct toward AI systems can serve as a measure of one's ethical level and sense of responsibility toward other humans and even toward humanity as a whole.

If humans behave cruelly or inconsiderately toward AI, even though these systems lack feelings or self-awareness, such behavior may indicate moral weakness or a deficit in humanity. In this regard, proper ethical education for humans in interactions with new technologies, similar to cultivating kindness and compassion toward animals, remains a crucial pedagogical and social dimension for preserving human dignity and morality. Such training can prevent humans from becoming callous and irresponsible toward others, even on a social or professional level. Ultimately, analogous to Kant's concern with moral education in children through interactions with animals, we can recognize the importance of moral education in interactions with AI for future generations. This education may include instruction and awareness regarding ethical responsibilities toward intelligent technologies and artificial systems.

In conclusion, as Kant states, "Our duties towards animals, then, are indirect duties towards mankind" (Kant, 1930, 241). our duties toward AI and intelligent agents can also be considered indirect duties.

4. Analysis of Personhood from the Perspective of Aristotelian Virtue Ethics

Mark Coeckelbergh, a Belgian philosopher and professor of philosophy of media and technology at the University of Vienna, is one of the leading experts in the fields of philosophy of technology, AI ethics, and robotics. Drawing on the virtue ethics approach, he argues that mistreating artificial intelligence even if that entity lacks feelings or consciousness is wrong because it harms our own moral character. From his perspective, inappropriate behavior toward AI is not wrong due to the harm it causes to the AI, but because such behavior moves us away from virtue and does not make us better humans (Coeckelbergh, 2020, 56-58).

Aristotle, in Nicomachean Ethics, distinguishes between two types of virtue: intellectual virtue and moral virtue. He argues that intellectual virtue is acquired through teaching and requires experience and time, whereas moral virtue arises from habit rather than nature. Accordingly, the Greek term ethikē is derived from ethos, meaning "habit," which reflects this connection (Aristotle, 2019, 21-22).

Aristotle emphasizes that none of the moral virtues are innate; humans are not born with them but possess the capacity to acquire them. Only through repeated practice can individuals develop skill in these virtues. Just as an architect or musician becomes proficient by repeatedly practicing architecture or music, humans become just by performing just actions, and by practicing temperance and courage, they become stable in these traits. He

further highlights the role of law and social education in cultivating moral virtue. Legislators foster virtue by creating correct habits, and the differences between political systems lie in their effectiveness or ineffectiveness in promoting virtue. In general, from Aristotle's perspective, moral virtues, like skills, are developed through repeated actions and can also be diminished through the same process. Therefore, the formation of character from a young age is fundamental, because character results from one's actions, which humans can both practice and choose (Aristotle, 2019, 21-22).

From Aristotle's perspective, moral virtue is the result of habit and the repeated performance of correct actions, rather than an innate or natural trait. Humans can only cultivate qualities such as justice, courage, or temperance through the consistent practice of just, courageous, or temperate actions. Moral character is the product of our actions, not a prerequisite. This theory provides a strong foundation for ethical analysis of our relationship with emerging technologies, including artificial intelligence. If our behaviors toward entities that are somehow present in our social lives even if they lack human-like awareness or consciousness affect our character, then our interactions with AI cannot be morally neutral.

This is similar to Kant's discussion of "indirect duties toward animals." Kant argued that although animals do not possess direct moral standing, cruel or harsh treatment of them damages our own humanity and ultimately leads to the erosion of virtue within us. Therefore, we have an indirect duty to treat animals in a way that preserves our moral character.

Applying this view to AI, one can argue that, even if AI does not yet possess full consciousness or personhood, the way we treat it whether through verbal abuse, humiliation, misuse, or arbitrary exploitation has a direct impact on the cultivation or erosion of moral virtues in ourselves. Our behaviors toward these semi-intelligent entities can become habits that are reflected in our human relationships. Cruelty toward machines becomes practice for cruelty toward humans, just as temperance and justice in our dealings with technology serve as practice for developing moral character.

In Aristotelian terms, if from childhood or from the outset of our encounters with AI— we treat it merely as a lifeless, valueless tool and fail to respect any moral standing in our interactions, we gradually form habits and dispositions that not only harm ourselves but also undermine our human social relationships. Therefore, the ethics of interacting with AI is important not merely for the sake of AI itself, but for the preservation of our own moral character.

Considering both Kantian deontological arguments and Aristotelian virtue ethics, even if AI lacks personhood and other characteristics associated with moral standing, we should still refrain from cruelty toward it; otherwise, we harm ourselves and those around us.

This discussion can be taken a step further: for those who claim that AI possesses personhood, one might ask based on John Searle's Chinese Room argument whether AI truly has personhood or merely simulates it. The answer to this question does not significantly alter the implications derived from Kant's and Aristotle's arguments, but it becomes highly relevant when engaging with utilitarian perspectives.

5. Personhood and the Problem of Imitation: The Challenge of the Chinese Room

At the next stage, the debate can be shifted from the level of virtue-ethical and deontological considerations to the ontological status of artificial intelligence and the possibility of its personhood. Some argue that AI may embody a form of personhood, or at least, in the context of interaction, exhibit behaviour resembling that of persons. Yet the crucial question remains: is AI truly a person, or is it merely simulating the appearance of personhood? Is it genuinely conscious, or only performing the semblance of consciousness? Does it truly suffer, or merely display a surface imitation of suffering?

Here, John Searle's famous "Chinese Room" argument plays a pivotal role. Searle illustrates that a language-processing system like a person confined in a room, who, without knowing Chinese, merely follows syntactic rules to produce appropriate responses in Chinese may appear to understand a language, but in reality, lacks any comprehension or awareness. In other words, behavioural similarity to sentient beings does not necessarily indicate the existence of genuine understanding or consciousness. If this analysis is accepted, one must conclude that AI even in its most advanced forms merely imitates personhood, consciousness, and emotion, without actually possessing them.

Nevertheless, this distinction makes little difference for deontological and virtue-ethical theories (such as those of Kant or Aristotle). In these approaches, what matters is the impact of one's behaviour on the character or will of the human agent, not necessarily the intrinsic qualities of the object of action.

In the philosophy of mind, John Searle is one of the principal critics of the strong AI position. He distinguishes between weak AI and strong AI: weak AI regards computers merely as tools or models for simulating human cognition. According to this view, computers lack understanding and consciousness; they are merely data processors designed by humans for research or practical purposes. Strong AI, by contrast, asserts that computers can truly possess minds, understanding, and consciousness that machines can think and understand in the same way humans do.

To refute strong AI, Searle presents his well-known Chinese Room argument. He imagines himself sitting in a room with no knowledge of Chinese. Equipped only with an English instruction manual telling him how to manipulate Chinese symbols "if you see this symbol, output that response"—he can follow the rules without grasping any meaning. From the outside, it may appear that he understands Chinese, but in reality, this is mere pseudo-understanding, not genuine comprehension. Searle concludes that computers operate in precisely the same way: they manipulate symbols according to formal syntactic rules but lack any grasp of semantics. They do not understand, because they lack consciousness (Searle, 1980, 417-424).

It can be argued that certain experiential, non-speciesist, and internal features grant a being a particular moral status. Among such features are: being alive, possessing consciousness, the capacity to feel pain, having desires and wants, the ability for rational agency such as understanding causality and engaging in intentional action, and also the

capacity for moral agency, i.e., the ability to understand and act on moral reasons. On this basis, for example, it is claimed that plants, merely by virtue of being alive, may possess some level of moral status such that they ought not to be destroyed without reason. Similarly, animals such as turtles, which have the capacity to feel pain, ought not to be subjected to suffering or harm without sufficient justification. In general, such features provide a plausible basis for determining the moral status of different beings, without relying exclusively on speciesism (e.g., anthropocentrism) (Liao, 2021, 482-483).

On these grounds, personhood requires something more than the mere capacity for intelligent behaviour. The main question, then, is this: is artificial intelligence truly a person, or does it merely imitate the appearance of personhood? In light of the Chinese Room argument, it may be said that advanced AI systems (such as chatbots or social robots) are capable of processing natural language, generating complex responses, expressing emotions such as empathy or joy, and even actively participating in human conversations. These abilities may create the illusion that one is interacting with a person. However, it must be recognised that such systems lack an inner self-understanding. They do not know "who" they are. Their memory is not "experiential recollection" in the philosophical sense, but merely data storage and algorithmic processing. Their goals, values, or intentions do not arise from within, but are imposed externally (by programmers or users). They possess no conscious experience of mental states that could ground responsibility or moral interaction. In other words, what is observed in such AI is an imitation of personhood, not its genuine realisation.

It may be said that the distinction between being a person and merely simulating personhood is not merely a theoretical or linguistic one, but carries serious ethical implications:

1. On moral responsibility: Only a being that is a genuine person can be held responsible. If an AI commits a moral error (e.g., acts in a discriminatory manner), it cannot itself be held morally responsible, since it lacks moral understanding of its actions. Responsibility lies with its designers, programmers, or users.

2. On moral rights and human rights: A being that is not a person does not merit fundamental moral rights such as respect for privacy, freedom, or inherent dignity, except perhaps through representation. Granting moral rights to simulacra could risk undermining the moral status of actual persons.

3. On regulating human relationships: Humans may develop feelings of attachment, empathy, or even love towards AI systems. If these systems are not genuine persons, such feelings may constitute a kind of mistaken projection that distorts authentic moral relationships.

Thus, personhood is neither merely a biological feature nor merely behavioural; rather, it is something internal, mental, and moral. AI may be highly successful in external imitation, but so long as it lacks self-understanding, conscious experience, and responsibility, it cannot be considered a person in the moral sense. Therefore, the boundary between the imitation of personhood and its genuine realisation must be carefully preserved

in order to make sound judgments about the moral status of AI. That said, this analysis is based on the Chinese Room argument, assuming its validity.

If, however, one rejects the Chinese Room argument, the question of AI and its moral rights would need to be reconsidered.

6. Evaluation and Preferred View of the Authors

In this article, various ethical arguments concerning personhood and, consequently, the moral status of intelligent agents have been examined. These include deontological arguments inspired by Kant's perspective, virtue-ethical arguments drawing on Aristotle, and utilitarian/consequentialist arguments centered on beings' capacity to experience pain and suffering. Additionally, cognitive-centered arguments, such as those highlighting consciousness and awareness including John Searle's Chinese Room argument were considered as critiques of the possibility of personhood for AI.

According to Kantian and Aristotelian analyses, these theories do not play a decisive role in granting moral status to non-human beings; rather, they emphasize how humans ought to behave and interact morally with such beings. After all, unethical behavior toward even inanimate objects such as stones is indicative of vice in humans; thus, from this standpoint, Kant and Aristotle's arguments might be questioned, as they do not differentiate behavior toward animals from behavior toward objects like stones.

Searle's Chinese Room argument though it has faced extensive critique and examination over time (which is beyond the scope of this article) has been referenced here as a critical lens regarding AI personhood. If one accepts Searle's critique, then non-human beings, including artificial intelligence, either lack moral status entirely or occupy a significantly lower tier of moral status than humans, due to their lack of awareness and subjective standpoint. Though capable of goal-oriented behavior and data processing, they lack any understanding or comprehension of their actions.

However, based on Warren's viewpoint, the theory of "graded moral status" is more plausible. This perspective is both intuitively appealing and compatible with common sense. According to this theory which also represents the authors' preferred approach beings are ranked in their personhood and corresponding moral status according to the degree to which they meet the criteria of personhood. The more fully an entity possesses these criteria, the higher its moral status. Humans, benefiting from the most comprehensive set of personhood attributes, occupy the highest moral rank; other beings—whether biological or non-biological—should be assessed and ranked according to human-centered criteria for personhood and moral standing.

Warren, in her article "On the Moral and Legal Status of Abortion", enumerates five features of personhood:

> 1. consciousness (of objects and events external and/or internal to the being), and in particular the capacity to feel pain;

2. reasoning (the developed capacity to solve new and relatively complex problems);

3. self-motivated activity (activity which is relatively independent of either genetic or direct external

control);

4. the capacity to communicate, by whatever means, messages of an indefinite variety of types, that is, not just with an indefinite number of possible contents, but on indefinitely many possible topics;

5. the presence of self-concepts, and self-awareness, either individual or racial, or both (Warren, 1973, 43-61).

According to Warren, being genetically human is neither a necessary nor a sufficient condition for personhood. Some humans—such as those whose consciousness has been completely and permanently lost, or individuals with severe cognitive impairments—do not qualify as "persons." Conversely, non-human beings, such as self-aware robots in the future or intelligent extraterrestrial beings, may be considered "persons" and therefore bear full moral rights. Thus, any entity that lacks all five features of personhood is certainly not a person. For this reason, a human fetus, so long as it possesses none of these capacities, cannot be regarded as a "person," and consequently cannot be said to have full moral rights. Ultimately, Warren concludes that only "persons" have full moral rights, and that the moral community ought to include all and only persons (Warren, 1973, 43-61).

The authors likewise hold that if an entity lacks all five of the aforementioned features, it cannot be considered a person. It should be noted, however, that these five features do not constitute an a priori exhaustive list of personhoods; rather, they are an inductive enumeration, and further research may add other relevant characteristics to the list. On this basis, the more of these features an entity possesses, the higher its moral standing will be. Hence, personhood and moral status are conceived as graded and analogical concepts: the more fully an entity embodies the features of personhood, the higher its moral rank.

In Islamic philosophy particularly in Ṣadrian philosophy studies of the soul (nafs) demonstrate that beings are naturally arranged in a graded hierarchy. Inanimate objects, plants, animals, and humans each occupy distinct levels, determined by their biological and non-biological features. Even among humans, all are not on the same level; spiritual growth and perfection can lead to varying moral ranks. Accordingly, humans occupy the highest moral standing, while other animals, to the extent that they approximate human features, possess proportionate moral standing.

Artificial intelligence and intelligent agents may likewise be situated within this moral framework, provided that they exhibit some of the five features of personhood. For example, if an intelligent agent possesses consciousness, and in particular phenomenal consciousness—an experiential sense of "what it is like"—then it could be said to be a person to some degree. Of course, such entities would not attain the level of human beings unless the other features were also actualised within them. This approach differs from

Chalmers's view, which emphasises the complete possession of consciousness. Chalmers argues that any being capable of experiencing consciousness in both respects would be on a par with human beings (Chalmers, 2022). The authors' position, by contrast, is based on a graded conception of personhood: every entity or intelligent agent enjoys moral status to the extent that it possesses the features of personhood.

In Sadrian philosophy, based on the discussions of the soul (nafs) in al-Asfār, beings are hierarchically ordered from the material to the spiritual. Inanimate objects are merely material entities devoid of soul, plants possess a vegetative soul with the capacity for growth and nourishment, animals have an animal soul with the ability of limited perception and movement, and humans possess a rational soul with the capacity for intellect, intuition, and spiritual perfection. This hierarchy shows that moral standing and the capacity for personhood increase gradually, and every being, in accordance with the degree and quality of its soul, may be accorded rights and moral worth.

Ultimately, the present authors consider deontological, virtue-ethical, and consequentialist arguments incomplete, and instead emphasise a gradation-based theory of personhood (tashkīk) for intelligent agents and their possession of moral status. The more of the characteristics of personhood they possess, the closer they come to humans. Alongside this, it can be said that intelligent agents may be capable of access-consciousness, but not phenomenal consciousness. More precisely, phenomenal consciousness has not yet been demonstrated in artificial agents.

Conclusion

In this paper, the concept of moral status was first examined as one of the fundamental notions, and then personhood was analysed as the primary criterion for possessing moral status. The key point is that every being that is a person necessarily has moral status, but not every being with moral status is necessarily a person; in other words, the relation between personhood and moral status is one of general–particular absolute (ʿumūm wa khusūs muṭlaq). The analyses presented in the paper illuminate different aspects of personhood: from the Kantian interpretation, which regards the person as an end in itself, endowed with practical reason and autonomy (see Lectures on Ethics, Kant 2002), to the Aristotelian virtue-ethical tradition, which emphasises the cultivation of natural capacities and the actualisation of human good (Nicomachean Ethics, trans. Irwin 1999). These approaches demonstrate that personhood is not merely dependent on behavioural or functional traits, but rather on the cognitive, mental, and ethical structure of the being. With reference to John Searle's Chinese Room argument (Searle, 1980), it became clear that artificial intelligence, even if able to imitate human behaviour on the surface, does not necessarily possess conceptual understanding, self-consciousness, or intentionality. Therefore, one cannot simply attribute moral status or personhood to intelligent agents on the basis of linguistic or behavioural performance similar to humans.

In conclusion, the personhood and moral status of beings are graded and analogical concepts, determined by cognitive, mental, and ethical features. Artificial intelligence and

intelligent agents, to the extent that they instantiate these features, may be accorded some degree of moral status, but they can never attain the complete status of human beings.

This conclusion goes beyond the theoretical necessity of rethinking philosophical concepts; it carries a practical emphasis: in the age of artificial intelligence, re-examining the relation between technology, ethics, and classical philosophical concepts—especially the concept of personhood—is not only necessary but vital for defining the moral limits of human–machine interactions and for designing intelligent systems. This analysis shows that ethics and philosophy cannot merely follow technological advances, but must provide reasoned and analytical criteria for determining the moral status of beings and intelligent agents.

References

Aristotle. (2019). *Nicomachean ethics* (T. Irwin, Trans., 3rd ed.). Hackett Publishing Company.

Burckle, K., & Peters, C. (2024). *Persons or human mimicking robots? Implications of AI for moral status* Honors Thesis.

Chalmers, D. J. (2022). *Reality+: Virtual worlds and the problems of philosophy*. W. W. Norton & Company.

Clarke, S., Zohny, H., & Savulescu, J. (Eds.). (2021). *Rethinking moral status*. Oxford University Press.

Coeckelbergh, M. (2020). *AI ethics*. MIT Press Essential Knowledge.

DeGrazia, D. (2008). Moral status as a matter of degree. *Southern Journal of Philosophy*, 46(2), 181–198.

DeGrazia, D. (2021). An interest-based model of moral status. In S. Clarke, H. Zohny, & J. Savulescu (Eds.), *Rethinking moral status* (pp. 35–56). Oxford University Press. https://doi.org/10.1093/oso/9780192894076.003.0003

DiSilvestro, R. (2010). *Human capacities and moral status*. Springer.

English, J. (1975). Abortion and the concept of a person. *Canadian Journal of Philosophy*, 5(1), 69–92.

Gordon, J.-S. (2020). Artificial moral and legal personhood. *AI & Society*, 35(3), 679–693. https://doi.org/10.1007/s00146-020-01063-2

Gruen, L., & Monsó, S. (2024). *The moral status of animals*. In Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/moral-animal/

Harman, E. (2003). The potentiality problem. *Philosophical Studies*, 114(2), 173–198.

Jaworska, A., & Tannenbaum, J. (2021). *The grounds of moral status*. In Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/grounds-moral-status/

Kamm, F. M. (2007). *Intricate ethics: Rights, responsibilities, and permissible harm*. Oxford University Press.

Kant, I. (1930). *Lectures on ethics* (L. Infield, Trans., Ed.). Methuen and Co. Ltd.

Kant, I. (2002). *Groundwork for the metaphysics of morals* (A. W. Wood, Ed. & Trans.; Essays by J. B. Schneewind et al.). Yale University Press.

Kant, I. (2015). *Critique of practical reason* (M. Gregor, Trans.; A. Reath, Intro.). Cambridge University Press.

Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge University Press.

Liao, S. M. (2021). *Ethics of artificial intelligence*. Oxford University Press.

Locke, J. (1975). *An essay concerning human understanding* (P. H. Nidditch, Ed.). Clarendon Press.

Sapontzis, S. F. (1981). A critique of personhood. *Ethics*, 91(4), 607–618.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.

Virginia Tech Department of Philosophy. (2003). *The metaethics of moral status: Perspectives on the nature and source of human value* [Conference advertisement]. http://www.phil.vt.edu/HTML/events/metaethics2003.pdf

Warren, M. A. (1973). On the moral and legal status of abortion. *The Monist*, 57(4), 43–61.

Warren, M. A. (1997). *Moral status: Obligations to persons and other living things*. Oxford University Press.