

Can a Robot Be a Person? De-Facing Personhood and Finding It Again with Levinas

Roberto Dell'Oro

*De ce terrible paysage,
Tel que jamais mortel n'en vit,
Ce matin encore l'image,
Vague et lointaine, me ravit.*

*Le sommeil est plein de miracles!
Par un caprice singulier
J'avais banni de ces spectacles
Le végétal irrégulier,*

*Et, peintre fier de mon génie,
Je savourais dans mon tableau
L'enivrante monotonie
Du métal, du marbre et de l'eau.
(Baudelaire)¹*

THE QUESTION “CAN A ROBOT BE A PERSON?” has emerged of late in the field of bioethics. It is a fitting provocation, an instigation to think impelled by technological advances, as have been the many ethical issues posed by developments in medicine over the past century.² Robots can do almost everything:

¹ “This morning I am still entranced - By the image, distant and dim - Of that awe-inspiring landscape - Such as no mortal ever saw. Sleep is full of miracles! Obeying a curious whim, I had banned from that spectacle - Irregular vegetation - And, painter proud of his genius - I savored in my picture - The delightful monotony- Of water, marble, and metal” (“Parisian Dream,” *The Flowers of Evil*, fleursdumal.org/poem/228).

² The emergence of robotics has triggered reflections at different levels, mostly with a concern for the epistemological and ethical dimensions of the impact of robots and artificial intelligence on human life. The anthropological reflection (in the sense of a *philosophical* anthropology) is less developed. A general introduction to the problem is found in Phil Husbands, “Robotics,” in *Cambridge Handbook of Artificial Intelligence*, ed. K. Frankish and W. M. Ramsey (Cambridge: Cambridge University Press, 2014), 269–95; Patrick Lin, Keith Abney, and George A. Bekey, eds., *Robot Ethics: The Ethical and Social Implications of Robotics* (London: MIT Press, 2012); Margaret

discharge complex logical operations, resolve algorithmic puzzles impossible to the average human brain, carry out operations on command, and even act *of their own accord*, posing the issue of whether their endowments, either in the cognitive field or in the sphere of autonomous decision-making, might not make them closer to us than we think.³ Thus the question: can a robot be a person?⁴

The question points to a doubt, a puzzlement about the nature of personhood with respect to its attributions. Our world has become the abode of *homo technologicus*, the impersonal space of the Neuter, as Levinas might put it: a world of objects exposed to the totalizing gaze of science, whose *outlook has become a doing*.⁵

The suggestions advanced by Japanese scientist Hiroshi Ishiguro are telling. For years, Ishiguro has worked toward the creation of interactive robots, specifically, robots with a human appearance or *androids*, on the premise that “we empirically know the effect of appearance is

A. Bode, *The Philosophy of Artificial Intelligence* (Oxford: Oxford University Press, 1990). In a critical vein, Hubert L. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason* (Cambridge, MA: MIT Press, 1994) and Guy Vallancien, *Homo Artificialis* (Paris: Michalon, 2017).

³ Japanese scientist Hiroshi Ishiguro, possibly one of the most renowned in the field, speaks of “human-robot symbiotic society.” See Hiroshi Ishiguro, “Studies on Interactive Humanoids,” in *Robo-Ethics: Humans, Machines, and Health*, ed. Vincenzo Paglia and Renzo Pegoraro (Rome: Pontifical Academy for Life, 2020), 67–102.

⁴ Although often robots, humanoids, and artificial intelligence are considered alike, this is not necessarily the case. To begin, one ought to distinguish between so called embodied and non-embodied machines. Furthermore, machines, whether embodied or non-embodied, may be provided with some form of artificial intelligence. They are either “stupid”—i.e., programmed to work automatically—or “intelligent”—that is, endowed with increasing cognitive and decisional abilities. Roberto Cingolani sums up the meaning of the distinctions in question thus: “The availability of increasingly powerful calculation machines is constantly extending the limitations of artificial intelligence. At the same time, it is allowing the development of increasingly performing embodied machines (provided with sight, touch, and biomechanical abilities), making realistic the assumption that robots are characterized by performances increasingly closer to those of the human” (Roberto Cingolani, “Robots and Intelligent/Autonomous Systems: Technology, Social Impact, and Open Issues,” in Pegoraro and Paglia, *Roboethics*, 34).

⁵ Though Levinas is critical of Heidegger, the latter has offered a powerful critique of technology (cybernetics) as the inevitable destiny of metaphysics. We dwell in the *Gestell*, in the “framework” provided by the totalizing outlook of technology and the forgetfulness of being. See Martin Heidegger, *The Question Concerning Technology and Other Essays*, trans. William Lovitt (New York: Harper, 1977), 14–17. Another interesting take on the relation of technology to modernity is Monette Vacquin, “The Monstrous as the Paradigm of Modernity? Or Frankenstein, Myth of the Birth of the Contemporary,” *Diogenes* 49/3, no. 195 (2002): 27–33. For a general philosophical analysis of the phenomenon of technology, Don Ihde, *Technology and the Lifeworld* (Bloomington: Indiana University Press, 1990).

as significant as behaviors in communication.”⁶ To tackle the problem of appearance and behavior, two approaches are necessary: one from robotics, and the other from cognitive science. The cross-disciplinary framework emerging from such interaction is *android science*, whose goal is the creation of robots with humanlike appearance, movement, behavior, and perception. Although further study is needed to address meta-level cognitive functions that more intelligent human-friendly robots might perform (intelligence embodiment, multi-modal integration, intention/desire, consciousness, and social relationships), the goal of scientific research, according to Ishiguro, is to “develop companion robots that can pass the Total Turing Test as a scientific and engineering goal.” The premise of the research in question is that the robot, having passed all the tests that evaluate its “total human-likeness,” will be “accepted as a member of our society.”⁷

The difficult question at stake here is “recognition.” The robot is not the person, but its human-like appearance and behavior allows the *other* to the robot, potentially a conversation partner, to be fooled into thinking it is a person.⁸ The relation between human and robot thus rests on a kind of game of pretense: though one knows the robot is not human, he/she can still deal with it as a social partner. This seems to be the hypothesis tested out experimentally by scientists like Ishiguro:

If a human consciousness recognizes the android as a human, he/she will deal with it as a social partner even if he/she consciously recognizes it as a robot. At that time, the mechanical difference is not significant; and the android can naturally interact and attend to human society. Verification of this hypothesis is not easy and will take a long time. However, it is an important challenge that contributes to developing deeper research approaches in both robotics and cognitive science.⁹

One may wonder what is the vision driving the research in question. The “human-robot symbiotic society,” what awaits us at the end of the experiment, is not a more *human* society, further humanized by the presence of robots, but the turning of humans into *inorganic*

⁶ Hiroshi Ishiguro, “Android Science: Toward a New Cross-Interdisciplinary Framework,” in *Robotics Research*, ed. S. Thrun, R. Brooks, and H. Durrant-Whyte (Berlin: Springer, 2007), 118.

⁷ Ishiguro, “Studies on Interactive Humanoids,” 72. The Total Turing Test (TTT) allows one to compare a robot manipulated by human operators and an autonomous robot controlled by developed technology.

⁸ *Geminoid*, a tele-operated android, can transfer the presence of the person to distant places. Tellingly, Ishiguro suggests that “through tele-operation, the operator—myself—could adapt to the Geminoid body and accept it as my own body” (“Studies on Interactive Humanoids,” 74).

⁹ Ishiguro, “Android Science,” 127.

intelligent life.¹⁰ The evolution at stake is thus somewhat reversed with respect to a teleological movement geared toward the human being. The trajectory rather entails a progressive liberation of the latter from any “flesh body” and the openness, made possible by technology, to a diversity of bodily forms that “may allow us to evolve further.” Evolution by technology pushes us beyond the limitations of life. “The ultimate aim of human evolution is immortality, achieved by replacing flesh and bones with inorganic material. Organic bodies are not a precondition for human existence in today’s world.... Humans come from, and return to, inorganic material. The human is currently almost a machine. We humans are going back to an inorganic state in the near future... we are trying to be an inorganic intelligent lifeform.... We can choose any kind of life forms.... That is, we can be released from the constraint of the human body.”¹¹

Can this world still be a world of persons? If so, what does it mean to retrieve a proper understanding of the person’s singularity in a spectacle of world-objects? Is not the attempt to expand the notion of personhood to include robots a function of the totalizing tendency in which everything becomes an object? The levelling of the difference between person and machine would then signal the failure to account for the subject’s separation from the world of objects, its *exteriority* to any totalizing pretense.

I take some of the notions elaborated by Levinas, whose echo is already evident in my *incipit*, as relevant to the question I am addressing in this paper. Mine will be less a direct confrontation with him, based on exegetical precision and punctual textual references, and more a personal appropriation of his mode of thinking, with which I find myself attuned.¹²

¹⁰ Ishiguro, “Studies on Interactive Humanoids.”

¹¹ Ishiguro, “Studies on Interactive Humanoids,” 94–100.

¹² Levinas has been a “companion in thinking” for many years. I have dedicated my STL thesis to Paul Ricoeur (“Antropologia ed etica nella *philosophie de la volonté* di Paul Ricoeur,” Pontificia Università Gregoriana, 1985) and my doctoral dissertation in theology to another phenomenologist, Dietrich von Hildebrand. See Roberto Dell’Oro, *Esperienza morale e persona: per una reinterpretazione dell’etica fenomenologica di Dietrich von Hildebrand* (Roma: Pontificia Università Gregoriana, 1996). Levinas has been an interlocutor for Ricoeur. On this, see Richard Cohen, *Ethics, Exegesis, and Philosophy: Interpretations after Levinas* (Cambridge: Cambridge University Press, 1992), 283–325. As for Dietrich von Hildebrand, the proximity to Levinas is evident at the level of their general sensibility, and this in spite of apparent reciprocal ignorance. Still, in a quasi-biographical narrative, concerning his encounter with phenomenology, and how he came to study with Edmund Husserl, Levinas mentions his friendship with Jean Héring, who was Husserl’s student in Göttingen, together with “an entire circle of young thinkers.” Since Hildebrand was part of it, Levinas must have known, at least, of the name. See Hans Reiner Sepp, ed., *Edmund Husserl und die Phänomenologische Bewegung. Zeugnisse in Text und Bild* (Freiburg/München: Karl Alber, 1988), 27–33. For a comparison between the two thinkers on their philosophy of love, see the introduction of John F. Crosby to Dietrich von

I begin with something like an archeological reconstruction of personhood in modernity, in order to locate the context out of which the question posed—"can a robot be a person?"—might take on meaning. Descartes, Hume, and Kant are the most important exponents of the story, their position emerging in direct contradiction to the classical metaphysics of the person, such as one finds in Thomas Aquinas. I see Levinas as having a complex relation with modernity, at once defined by a positive retrieval of Descartes and Kant, and critical of the anthropological dualism effected by the Cartesian *cogito*.¹³ Levinas rejects the rationalist perspective of a *bodiless mind*, a person reduced to her cognitive capacities, no less than the empirical version of a *mindless body*, a person reduced to the external stimuli of sensations and impressions registered by the mind.

Contemporary bioethics, however, especially in the Anglo-American version of it, is mostly defined by such an understanding of personhood: the person does not "come to mindfulness" out of its bodily conditions, nor does she persist, when no longer conscious, in bodily presence. According to Peter Singer, one of the main voices in bioethics, personhood is transitional: it passes from being to being like a "thing," as long as certain dimensions of actual empirical consciousness are present. Thus, a dolphin, a chimpanzee, a higher mammal, perhaps a robot can be a person.¹⁴

On the other hand, as Levinas suggests, to be a person is to be "manifested in the exteriority of the face, which is not the disclosure of an impersonal Neuter, but *expression*, that is, the presence of an infinite idea that always exceeds the idea of the other in me."¹⁵ If so, a robot cannot be a person. In what follows I try to say why I think this is the case.

Hildebrand's *The Nature of Love* (South Bend: St. Augustine's Press, 2009), xxxi. More broadly, the interesting article of Alexander Montes, "Toward the Name of the Other: A Hildebrandian Approach to Levinasian Alterity," *Questiones Disputatae* 10, no. 1 (2019): 82–109.

¹³ More than on the various dualisms asserted by Descartes, Levinas focuses on the latter's idea of the infinite, as providing the point of entry into the meaning of the relation between same and other: "This relation of the same with the other, where the transcendence of the relation does not cut the bonds a relation implies, yet where these bonds do not unite the same and other into a Whole, is in fact fixed in the situation described by Descartes in which the 'I think' maintains with the Infinite it can nowise contain and from which it is separated a relation called 'idea of infinity'" (Emmanuel Levinas, *Totality and Infinity: An Essay on Exteriority*, trans. Alphonso Lingis [Pittsburgh: Duquesne University Press, 1969], 48).

¹⁴ For a survey of various positions on personhood in bioethics, see the chapter on moral status in Tom L. Beauchamp and James F. Childress, *Principles of Biomedical Ethics*, 7th ed. (New York: Oxford University Press, 2013), 62–100.

¹⁵ Levinas, *Totality and Infinity*, 50–51.

PERSON: A HISTORICAL RECONSTRUCTION*Facing the World in Wonder: Thomas Aquinas and the Person as Spirit Incarnate*

For Aquinas, who comments on Aristotle's metaphysics, philosophy begins in *wonder*.¹⁶ Wonder is the attitude that throws us back onto ourselves, in stunned astonishment at the sheer being there of things. Later on, such astonishment will engender perplexity about the meaning of things. Prior to the activity of questioning perplexity or doubt, wonder entails a kind of trust, a confidence (*fides*) in the natural goodness of reality, a confidence which is a love: being is promising and good (Genesis says: "God saw that it was very good"). For Thomas, who inherits the insights of the entire Christian metaphysical tradition, being is the miracle of gratuitous generosity, a "being there" without explanation, out of a source that gives. Creation is a gift of the Origin whose "coming into being" remains *in excess of* the ontological relations that define being in its "becoming," such as form and matter, formal and final causality, potentiality and actuality, etc. A being that is becoming already presupposes its "being given into existence," now charged with the promise of further development.¹⁷

But what does it mean "to be"? Thomas refers to two aspects of being: the "in-itself aspect" of being (*substance*), and its "towards-others aspect" (*relationality*). "To be" is "to exist," and to exist is to be *an-integrity-in-relation*. One might say that existence is the actualization of the energy of being, the gift from a source that offers itself, and whose communicative aspect is being participated to each existent.¹⁸ Consider the following quotation from Gerald Phelan:

The act of existence (*esse*) is not a state, it is an act, the act of all acts, and therefore must be understood as act and not as a static definable object of conception. *Esse* is dynamic impulse, energy, act—the first,

¹⁶ The statement is originally from Plato, who lets Socrates say: "I see, my dear Theaetetus, that Theodorus had a true insight into your nature when he said that you were a philosopher; for wonder is the feeling of a philosopher, and *philosophy begins in wonder*," *Theaetetus* 155c-d, in *The Dialogues of Plato*, trans. B. Jowett (New York: Random House, 1937), vol. 1, 157. For Aristotle, "It is owing to their wonder that men both now begin and at first began to philosophize," *Metaphysics* 982b, in *The Basic Works of Aristotle*, trans. Richard McKeon (New York: Random House, 1941), 692. According to Thomas Aquinas, wonder (*admiratio*) is "a kind of desire (*desiderium*) for knowledge; a desire which comes to man when he sees an effect of which the cause either is unknown to him, or surpasses his knowledge or faculty of understanding" (*Summa theologiae* I-II, q. 32, a. 8).

¹⁷ The third proof, so called "on possibility and necessity," speaks of the contingency of being. See Thomas Aquinas, *Summa theologiae*, I, q. 2, a. 3.

¹⁸ "*Being* means that-which-has-existence-in-act.... Now any designated form is understood to exist actually only in virtue of the fact that it is held to *be*.... It is evident, therefore, that what I call *esse* is the actuality of all acts," in *An Introduction to the Metaphysics of St. Thomas Aquinas*, ed. James F. Anderson (South Bend: Regnery/Gateway, 1953), 22.

the most persistent and enduring of all dynamisms, all energies, all acts. In all things on earth, the act of being (*esse*) is the consubstantial urge of nature, a restless, striving force, carrying each being (*ens*) forward, from within the depths of its own reality to its full self-achievement.¹⁹

Of course, one cannot fully grasp all this without reference to its theological underpinning. Thomas talks about being, but ultimately thinks about God, the Christian God of creation.²⁰ In this conceptual framework, the person also finds her place. Unlike other beings, the person is not just being-in-itself but being-coming-to-itself in self-presence (Thomas speaks of *reditio*, return unto itself).²¹ The person is a mindfulness of being in its totality, a complete openness to its fullness and, thus, most fully being. Indeed, the person is the most perfect of substances, because of her ecstatic openness to the totality of being, an openness actualized in knowledge and will.²² Such openness is not just a function of one particular aspect of the person, say the mind. It is the function of the entire being that is the person in the unity of its principle—i.e., its *soul* (for Thomas, like Aristotle: *anima est quodammodo omnia*)—because the soul is the form of the body (*anima forma corporis*).²³

This is an important point: when thinking about the person, Thomas always points to the unity of body and soul. In encountering

¹⁹ Gerald Phelan, "The Existentialism of St. Thomas," *Selected Papers* (Toronto: Pontifical Institute of Medieval Studies, 1967), 77. Quoted in Norris W. Clarke, *Person and Being* (Milwaukee: Marquette University Press, 1993), 9.

²⁰ To my knowledge, there is no reference to Thomas Aquinas in Levinas. The question of whether the former might not fall into Heidegger's general condemnation for the Western metaphysical tradition as "onto-theological" might have influenced Levinas' own way of reading the tradition. Jean-Luc Marion, a student of Levinas, is keen in rescuing Thomas Aquinas from the accusation. Consider the following: "To think *esse* starting from God, but not in inverse order (in the way of *metaphysica* and of Heidegger as well), allows Thomas Aquinas to free the divine *esse* from its—tangentially univocal—comprehension starting from what philosophy understands by being, entity being of the entity, in a word to mark the distance—an "infinitely infinite distance"—from the creature to God (Pascal). . . . One could say that such (divine) *esse* keeps within itself the transcendence that opposes the act of being to the *esse commune* of entities' . . . Therefore, God without being (at least without *this* being) could become a Thomistic thesis." The quotations are from the essay titled "Thomas Aquinas and Onto-theo-logy," in Jean-Luc Marion, *The Essential Writings*, ed. Kevin Hart (New York: Fordham University Press, 2013), 306–307.

²¹ "*Illa quae sunt perfectissima in entibus, ut substantiae intellectualis, redeunt ad essentiam suam reditione completa*," Thomas Aquinas, *De Veritate*, q. 1, a. 9.

²² The point is central in the reinterpretation of Rahner and so-called "transcendental Thomism." See Karl Rahner, *Spirit in the World*, trans. William Dych (New York: Continuum, 1994).

²³ For a systematic analysis of Thomas Aquinas's anthropology, see the classic work of Sofia Vanni Rovighi, *L'antropologia filosofica di San Tommaso D'Aquino* (Milano: Vita e Pensiero, 1965).

the body, one encounters the person in her totality, in her spiritual presence. The person is, in this sense, “spirit incarnate,” as Rahner would say, an integrity of being, a substantial unity. In this integrity of being the person is also infinite openness, infinite intentionality: *homo capax Dei*.

One might summarize by saying that for Thomas Aquinas the person entails three dimensions: to be an existent, that is, an integrity of being unto itself (incommunicable substance); actively open to other being (substance-in-relation); and passively receptive to the totality of being (receptivity). Incommunicability, relationality, and receptivity point to a balance between passivity and activity, *passio* and *conatus*, in the person.²⁴

Here is where also the question of *potentiality* in the person comes into play. To be a person as substance-in-relation means that the person (like being itself) is always in a process of becoming: openness and receptivity are the conditions for growth and dynamic passing from potency to act. To be a person is always to be in-potency, in a process of progressive actualization toward a more perfect fulfillment. To be a person is to be an energy of transcendence.

The Thomistic synthesis will undergo a shift of paradigm as it moves into modernity, with important changes in the philosophical and theological disposition of the late Middle Ages, the age of nominalism. First, with respect to the concept of God: the stress is now on God’s transcendence, on his distance from the world, as well as on his will, which is inscrutable and absolute—unlike Thomas, who sees the eternal law as the height of rationality. There are also changes in the concept of nature: both the distance of God and the inscrutability of his will leave the world devoid of signs of his presence. The world becomes “de-sacralized” or *secularized*, but, because of this, also available for the exploration of man, his engagement with things, and the general project of natural discovery. Nature can no longer hide—contra Heraclitus, for whom “nature loves to hide.”²⁵

God is beyond nature, infinitely powerful, and inaccessible. Now stripped of all the signs of divine communication, being is no longer conceived as communicative (*actus essendi*, act of existence), the energy of relationality, but as a being-unto-itself, distant and cold. A progressive impoverishment in the understanding of being sets in.²⁶

²⁴ Clarke, *Person and Being*, 25–110.

²⁵ *Physis kruptesthai philei*. The fragment belongs to Heraclitus’s *ipsissima verba* in the Diels-Kranz collection (B123). See Jonathan Barnes, *Early Greek Philosophy* (London: Penguin, 1987), 122.

²⁶ A thorough historical analysis can be found in Emerich Coreth, *Metaphysik: Eine Methodisch-Systematische Grundlegung* (Innsbruck: Tyrolia, 1980), especially 15–47. The impoverishment in the conception of being, following the nominalistic “effectual history,” and in critical dialogue with Heidegger, has been argued by Gustav Siewerth, *Das Schicksal der Metaphysik von Thomas zu Heidegger* (Einsiedeln:

Consider as an example the loss of an analogical understanding of reality and the univocal reduction to particularity, taking place especially with Duns Scotus.²⁷

One aspect of the change is paramount: the emphasis on *essence* rather than *existence*, and the subsequent loss of the notion of being as energy of communication. Being is no longer seen in the fullness of its over-determinacy, in its plenitude (Jacques Maritain speaks of the “generosity of being”), but as a thing-like presence, retracted unto itself, an object facing a subject. The essentialization in question signals the end of a teleological understanding of nature, the denial of its spontaneous finality. Nature knows no intrinsic destination. It is not a dynamic system that gives itself, as in the etymology of *physis*, but a static object, to be grasped by a subject.

THE PERSON AS SUBJECT AND THE DUALISM OF RENÉ DESCARTES

The story of modernity begins with a skeptical puzzlement about the nature of reality, the place of God in it, and the destination of the self. The modern *Denkform* is new with respect to its historical predecessors. If the classical tradition of metaphysics, from Plato to Aristotle and Aquinas, begins in wonder, Descartes’s philosophical discovery, *cogito ergo sum*, emerges in the wake of doubt, and functions as the condition of possibility for the reconstruction of an image of the world now certain, finally resting on secure foundations beyond any skeptical assault.²⁸ There are losses and acquisitions following such a shift of paradigms.

The losses: Descartes no longer understands being in its character of act, energy, power, communicative to mind. Furthermore, being is no longer endowed with intrinsic value; as such, it was able to speak to the mind. Now a reversal of directionality in the relation between mind and being sets in: the mind dictates to being the conditions of its meaningfulness. “I think,” that is, the epistemological certainty, precedes and grounds “I am,” that is, the ontological constitution.²⁹

Johannes, 1959). See also Etienne Gilson, *Being and Some Philosophers* (Toronto: Pontifical Institute of Medieval Studies, 1952).

²⁷ “Being is said in many senses” (*to on legetai pollakos*), says Aristotle in *Metaphysics*, Book IV, 2 (*The Basic Works of Aristotle*, 732). Similarly for Thomas, *ens multiplicitur dicitur*.

²⁸ The reference is to the *Discourse on the Method*, Part IV, but the statement recurs in the *Second Meditation* as well. See *The Philosophical Works of Descartes*, trans. Elizabeth S. Haldane and G. R. T. Ross, vol. 1 (Cambridge: Cambridge University Press, 1981), 102.

²⁹ Marion has highlighted the ambivalence of Descartes’s metaphysics, especially with respect to the idea of God. His focus is on Descartes’s doctrine of “eternal truths,” which cannot be ascribed both to the world and to God. Thus, all theology, understood as meaningful talk about God, is no more than a blank space or a white page in Descartes’s writings. See Jean-Luc Marion, *Sur la théologie blanche de Descartes*.

I said there are also important theoretical acquisitions in this shift. The most important is the turn to subjectivity, an “anthropological turning point” which, later on, Kant will compare to a Copernican revolution: to be a person is to be a subject. In this revolution, the language of substance gives way to the language of subjectivity, self-reflexivity and interiority. The person is not just a substance, even if the most perfect, in ontological continuity with the others. It is an entirely different being, discontinuous with the rest of creation, because of its ability to think. Pascal will say: “Man is only a reed, the most feeble thing in nature, but a thinking reed.”³⁰

To the subjectification of substance follows the objectification of being: being becomes an object for a mind that is a subject. In its neutrality, being is simply a raw reserve of resources available for human exploitation.³¹ Without a *telos*, an end (final cause), nature cannot account for its meaningful origin (*arche*). Nature becomes a neutral thereness, stripped of value, without formal or final cause.³² It is the realm of effective causality, a network of forces linked together by mechanic interaction.

There is no denying that the main conquests of science rest upon such an understanding of reality. In his reconstruction of the entire trajectory of modern thought, especially in terms of its scientific advances, Edmund Husserl offers the following account:

The exclusiveness with which the total world-view of modern man lets itself be determined by the positive sciences and be blinded by the “prosperity” they produced, meant an indifferent turning away from the questions which are decisive for genuine humanity. Fact-minded science excludes in principle precisely the questions which man finds the most burning: questions of the meaning or meaninglessness of the whole of human existence.³³

In such a *Weltanschauung*, questions of meaning will, subsequently, be bracketed as unimportant—in fact, be expunged from the epistemic drive toward verifiability. What then remains for the

Analogie, création des vérités éternelles et fondement (Paris: Presses Universitaires de France, 1981).

³⁰ Blaise Pascal, *Pensées*, trans. A. J. Krailsheimer (London: Penguin, 1995), 347.

³¹ William Desmond insists on such dialectic of subjectification and objectification in modernity. See his *Ethics and the Between* (Albany: SUNY Press, 2001), 17–47; and *God and the Between* (Oxford: Blackwell, 2008), 17–45.

³² The modern gaze is akin to an act of unveiling, and the forcing of nature (Galileo) signals the end of teleology. So Spinoza, for whom “nature has no fixed goal and all final causes are but figments of human imagination” (*naturam finem nullum sibi praefixum habere, et omnes causas finales nihil nisi humana esse figmenta*), Baruch Spinoza, *Ethics*, trans. Samuel Shirley (Indianapolis: Hackett, 1992), 59.

³³ Edmund Husserl, *The Crisis of European Sciences and Transcendental Phenomenology*, trans. David Carr (Evanston: Northwestern University Press, 1970), 5–6.

subject, when facing such a mechanized understanding of reality? It sees itself as other to the mechanical world or at least as irreducible to it. The world is purposeless, but the subject is purposeful, active with respect to being, not passive. It will provide being with the value being does not possess intrinsically. With Kant, the subject turns into "a self-assertive subjectivity." The person becomes the source of value, a noumenal being endowed with infinite value, with dignity.³⁴

The separation of mind and being at the ontological level effects important changes in the understanding of the person. For the person too will now be defined by an intrinsic split, a separation between body and soul. The person is no longer grasped in the unity of a single substance but dissolved in the dualism of two separate substances. From the idea of the human being as a substantial unity of body and soul, as incarnate being, a separation sets in between soul (mind) and body. The soul and the body will be seen as separated from one another, with the soul losing its original meaning of principle of life and unity. Furthermore, the soul progressively becomes intellectualized as "mind," and this in opposition to the body.

The separation of mind and body opens up a dualism and a reduction in the understanding of the person: the person is her mind, independent of her body. The separation also comes with a new attribution of value: the mind is higher than the body. Here Descartes inherits the dualistic theological logic of modern devotionalism (*devotio moderna*), with its emphasis on the denial of the body.

What emerges is a kind of deracinated person, a personhood without body, a personhood reduced to its cognitive faculties. To be a person no longer means to grow into the biological space defined by a specific life-principle (the soul), but to be able to actualize the faculties that are proper to the traits of a "thinking substance." The person will be such only because *cogitans*, a capacity to think; and this independently of the body, whose connection with the mind is now viewed as entirely accidental. The human body, now separated from the spiritual principle, becomes like a machine (*res extensa*), which entails the loss of the incarnate self, and the intellectualization of the spirit. As the "other" substance in the human composite, the body is left open to the manipulative intervention of the superior principle. The emergence of anatomy in medicine focuses on the body as inert entity (*Körper*), rather than lived-reality (*Leib*).

We need to understand the profound implications of the dualism in question. We are inheritors of such dualism, and I will say that to ask about the personhood of a robot is to fall *de facto* into the trap set by such pre-comprehension. One can look at the consequences of the dualism in question from two angles: either the angle of a *bodiless mind* or the angle of a *mindless body*. We have here the two developments

³⁴ Desmond, *Ethics and the Between*, 17–47.

that follow Cartesian dualism: rationalism and empiricism. Let me begin with the latter.

THE DECONSTRUCTED SELF OF DAVID HUME

The empirical line, pursued by Hume and the British empirical tradition, develops all the way to nineteenth century utilitarianism. The body receives sensations, sensible impressions from experience and, progressively, the mind builds a sense of identity out of the congeries of such impressions.

What is the person? It is the product, the net result of external stimuli registered by the mind. One can see that because such impressions are seen in their *transitional* capacity to impress the mind, the identity of the person will only be the result of the *psychological*, rather than substantial, ability of the mind to retain impressions through memory. For Hume, the person is not an integral center of being (substance), but the flow of impressions that come and go, insofar as they are retained by memory. Thus, the paradox: there is "person" only insofar as there is actual consciousness of sensations and impressions or, as we might say today, empirical stimuli.³⁵

Consider the thought experiments of contemporary bioethicists: when a person loses her consciousness, does she become a different person? Can we have two persons when a patient loses the ability to reason or to retain memories, etc.? The case becomes particularly acute when it comes to Alzheimer patients.³⁶

For Hume, one of the great problems was the experience of sleep: when I am asleep, do I cease to be the person I was? Am I waking up every time a different person? Hume looked into the mind in order to find a unity to the flow of impressions and found nothing. He concluded that *there is no person* as a substantial principle of integrity-in-communication.

If there is a person only insofar as there is *actual* consciousness of sense impressions, then the body is no longer central in providing the condition for the mind's continuity, nor is the embodied-self understood as dynamically growing into its full potential. Consciousness of sense impressions means that both the preconditions of bodily development and the fading of mental functions in a still operative body will not affect the presence of personhood. The person does not "come

³⁵ In the *Treatise* (Book I, Part IV, Section VI), David Hume speaks of the self as "nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in perpetual flux and movement" (*A Treatise of Human Nature*, ed. L. A. Selby-Bigge [Oxford: Clarendon, 1988], 252).

³⁶ For an early example of the literature on this, see Allen Buchanan, "Advance Directives and the Personal Identity Problem," *Philosophy and Public Affairs* 17 (1988): 277–302. See also Helga Kuhse, "Some Reflections on the Problem of Advance Directives. Personhood and Personal Identity," *Kennedy Institute of Ethics Journal* 9, no. 4 (1999): 347–64.

to mindfulness" out of its bodily conditions, nor does she persist in its no longer conscious bodily presence. Prenatal life will not be personal. Nor will the life of the senile demented be the life of a person.

Two final points, one relative to freedom of the will, the other to the general conception of morality as based on emotions, rather than reason. For Hume there is no freedom of the will because the person, as he understands it, is entirely determined. Freedom of the will is a mask, a deception, at best the epiphenomenon of something else to which it can be reduced. Humean determinism will cast long shadows, all the way to contemporary philosophies of mind, in which the de-personalization of the human being is complete.

A second point: the deconstruction of personhood and the denial of freedom rest upon a conception of morality based on emotions, rather than reason. The feeling of sympathy towards other human beings will become the basis of morality. But how fleeting such a feeling is! The Humean retrieval of emotions, feelings and, in general, of the affective dimensions of the person, is *de facto* equivocal: it stands in the wake of Cartesian dualism and cannot provide a secure basis for moral judgment. Hume will become the father of non-cognitivism.³⁷

THE PERSON AS UNIVERSAL EGO AND AUTONOMOUS SUBJECTIVITY: IMMANUEL KANT

I come to the rationalist line of development, following Cartesian dualism. Kant saw that the empirical ego is parasitical on a more original, *a priori*, notion of the self. The transcendental ego is the condition of all ordered experience *qua* experience of a unitary I.³⁸ Kant's philosophical anthropology represents a reaction to the fragmented self of Hume: the universal/transcendental ego is the formal capacity to gather the multiplicity, the manifold into an ordered unity.³⁹ With Kant, we have a more rigorous understanding of the meaning of personhood which, however, still stands within the "effectual history" of Cartesian dualism. Such dualism will be rendered by Kant in terms of a gulf between the phenomenal and the noumenal sphere of reality. To

³⁷ At the same time, the rehabilitation of the emotional sphere in Hume has not remained without important consequences, for example, in certain strands of contemporary feminist ethics. See Annette Baier, "Hume, the Women's Moral Theorist?," in *Women and Moral Theory*, ed. E. Feder Kittay and D. T. Meyers (Totowa, NJ: Rowman & Littlefield, 1987) and Alisa Carse, "The Voice of Care: Implications for Bio-ethical Education," *Journal of Medicine and Philosophy* 16, no. 1 (1991): 5–28.

³⁸ See Kant, *Critique of Pure Reason* on the "transcendental deduction."

³⁹ For the empiricist, the manifold in question can be brought to unity in a derived synthesis which, however, can easily fall apart without an *intrinsic* principle of unity. Such a principle points to a prior and not just derived synthesis, i.e., the "synthetic a priori": "The empiricist fails to see that a derived synthesis is possible only on condition of a prior synthesis, which is, in this respect, the condition of its possibility," William Desmond, *Desire, Dialectic, & Otherness: An Essay on Origins*, 2nd ed. (Eugene: Cascade, 2014), 69.

be a person is to belong to this noumenal realm, a realm of rationality and capacity for absoluteness. It is a realm of freedom, rather than natural necessity, such as the one entailed by Newtonian physics.

The noumenal realm also is a realm of value in a world stripped of it, as if the value sucked out of the objective world is now being channeled back into the subject, who becomes a being of infinite value. Question: where does the value of the subject come from if the world is valueless? How can something like a being *endowed with value* emerge from such a valueless world?

With Kant, the notion of personhood is clearly recognized in its moral significance. To be a person is to be a moral being able to exercise moral agency. In this resides the dignity of the person, its infinite value. To be a person is to be *autonomous*—i.e., to grant meaning, not to receive it (from God, religion, nature, etc.).

After Kant, the story of autonomy is the story of its progressive radicalization. There is a logical trajectory that runs from Kant to Nietzsche, from the person as power of autonomous decision making to the person as will to power.⁴⁰ In a world devoid of intrinsic value, and with an assertive (autonomous) notion of the person, the human also becomes an object. The objectification of the world turns into the objectification of the human. Having mechanized the world, we then end up mechanizing ourselves, looking at ourselves as mechanical systems, machines. The machine is the perfect expression of the modern outlook on reality: it is entirely constructed and available, that is, disposable. It is the perfect “object.”

We too become like constructed mechanisms. We cry freedom of choice and autonomy but fall victim to all sorts of constructionisms: social (Marxism), psychological (Freud), neurological. Think of the attempt of cognitive science to reduce the human mind to the physiological circuitry of the brain, and explain away consciousness in terms of deterministic happening, entirely reducible to neurological functionality.

The essence of the machine: to be entirely passive to our own construction. In the case of the robot, its apparent activity, whether cognitive or practical, is entirely determined by us. Even when robots think for themselves, they do not really think. They only carry out pre-ordained programs based on algorithmic laws of our own devise.

TO BE A PERSON: A PHILOSOPHICAL ACCOUNT

In light of this historical reconstruction, I now want to offer a more systematic understanding of the person. I begin with a provisional definition, something like a heuristic statement, aware of the fact that

⁴⁰ For this interpretation already, see Henri De Lubac, *Le drame de l'humanisme athée* (Paris: Cerf, 1999); and Romano Guardini, *Das Ende der Neuzeit. Ein Versuch zur Orientierung* (Würzburg: Echter, 1951).

Levinas would contest the possibility to “de-fine” a person. As a manifestation of the Infinite, the person escapes definition, at least in the sense that definitions are able to circumscribe objects. No objective definition of the person is possible without passing through a “lived subjectivity,” the *event in being* Levinas calls “psychism”: this is a *way of being* resistant to totality, in this case the one pursued by the search for definitional boundaries. Perhaps a better way to pose the question is not to enquire “what is a person?” but rather “who is a person?”

I think Levinas would push us to think of the person in terms of *an incarnate singularity, coming to itself, in openness to the Other*. I will parse out these three elements, incarnate singularity, selfhood, and openness to the Other, in order to arrive at an understanding of the person that owes greatly to Levinas, even if in the end it will differ from his in significant ways.

THE PERSON AS AN INCARNATE SINGULARITY

To be a person is to be individuated, to be an individual, to be one with oneself (already Boethius and Thomas Aquinas speak of *individua substantia*). Such individuality is expressive of a certain incommunicability (Roman law defines the person as *sui iuris et alteri incommunicabilis*). We are singular beings, though similar to others (brothers, monozygotic twins, or even clones). Levinas stresses this sense of incommunicability, independence, and separation, in terms of a break from any notion of participation:

One can call atheism this separation so complete that the separated being maintains itself in existence all by itself, without participating in the Being from which it is separated—eventually capable of adhering to it by belief—One lives outside of God, at home with oneself; one is an I, an egoism. The soul, the dimension of the psychic, being an accomplishment of separation, is naturally atheist.⁴¹

The singularity in question is not the result of a statement of singularity; it is not a vindication of singularity (as “constituted,” it would already belong to “the order of thought”). It is not constructed, but *given in the flesh*, in the body that we have, the “body that we are,” to echo Gabriel Marcel. This singularity is incarnate. Only a human being can be a person—i.e., someone who belongs to the human species—because the singularity in question is embodied in the singularity of a human flesh.⁴² Consider the following quotations from *Totality and Infinity*:

⁴¹ Levinas, *Totality and Infinity*, 58.

⁴² The question of the attribution of the notion of person to spiritual creatures, such as angels, or to God, would take us too far afield and cannot be addressed here. The focus remains on the human person. In critical distance from Heidegger, for whom

The sensibility we are describing starting with enjoyment of the element does not belong to the order of thought but to that of sentiment, that is, the affectivity wherein the egoism of the I pulsates.⁴³

Sensibility established a relation with a pure quality without support, with the element. Sensibility is enjoyment. The sensitive being, the body, concretizes this way of being, which consists in finding a condition in what, in other respects, can appear as an object of thought, as simply constituted.⁴⁴

This is relevant in addressing the question of double personality raised in the wake of the empiricist understanding. To say “I” (as we will see later, when talking about the self) is to actualize an energy of being that might still be dormant in the body, but that in its elementality reminds us of our being given to be *in the body that we are*.

To be a person is to be in a potential, constant state of growth. The person is always the promise of something more. We could say that such elementality cannot yet be objectified; it is pre-objective, in the sense of being felt, rather than determined in itself, as an object.

We feel ourselves, first of all, in bodily immediacy. If the dimension of self-insistence is prevalent at this point, it is a self-insistence that is also already a community with others. Levinas brings attention to the phenomenon of the *face*. In my face I say, “I am”; this always also is a “here I am,” that is, “I am with.” One must stress the flow-like, rather than fixed, character of this incarnate singularity.⁴⁵ The singularity in question is also a site of flow and passage, undergoing the world. The incarnate singularity is a passion of being.⁴⁶

sensibility means the reversal of praxis over theory, Levinas stresses a more elemental notion of sensibility, which he calls “enjoyment” (*jouissance*): “Sensibility, Levinas discovers, does not first emerge as praxis caught up in the larger network of “in-order-to” (*das Um-zu*)—the “referential totality” (*Verweisungsganzheit*)—which ultimately implies Dasein. Rather, sensibility is first the sheer enjoyment of sensations, a ‘care-free’ contentment with sensing itself. Embodiment, sensibility, flesh, is, first a self-satisfaction and an enjoyment of elemental sensations, the sun on one’s arms, the breeze in the air, indifferent to the higher-level significations of instrumentality and theory,” Richard Cohen, *Ethics, Exegesis and Philosophy*, 154.

⁴³ Levinas, *Totality and Infinity*, 135.

⁴⁴ Levinas, *Totality and Infinity*, 136.

⁴⁵ Here is where a metaphysics of substance falters.

⁴⁶ One will note the difference between this elemental, suffering being (*passio*) and the self-positing ego of transcendental philosophy (*conatus*). Leibniz had a premonition of the meaning of this suffering being when he distinguishes between “perception” and “apperception”: the self as flesh is perception not yet conscious in the distinct sense of apperception.

THE SELFHOOD OF THE HUMAN PERSON

The person is a being that comes to itself in self-awareness and reflectivity, in action and the development of moral agency. The self is not just given to itself. It becomes a self, it becomes a subject, indeed, a "thinking self" (Descartes). This coming to self, as Hegel points out, can only be possible in the intermediation with what is other. The third dimension in the definition provided above, that of "openness to the Other," is the last only in a temporal, rather than logical sense: it already subtends the other two components. The relation to the other is "older" than the relation to the self: "One may legitimately ask oneself whether the internal discourse of the *cogito* is not already a derivative mode of the conversation with the other; whether the linguistic symbolism that the soul uses in 'conversing with itself' does not suppose a dialogue with an interlocutor other than itself; whether the very interruption of the spontaneous impulse of thought reflecting upon itself, all the way down to the dialectical alterations of reasoning where my thought separates from and rejoins itself as if it were other than itself—whether this interruption does not bear witness to an *original and foregoing* dialogue."⁴⁷

I said before that the incarnate singularity already undergoes the world. The body is always a medium of exchange, it is never only "mine": "The subjectivity of the subject, its very psyche, is a possibility of inspiration. It is the possibility of being the author of what has been breathed in unbeknownst to me, of having received, one knows not from where, that of which I am the author. In the responsibility for the other we are at the heart of the ambiguity of inspiration."⁴⁸

FIRST ETHICAL SELVING: INTENTIONALITY

We come to ourselves in knowledge and action. The first component speaks to the exercise of mindfulness as an intentional act, a communion with what is other to us, an "object" toward which we move (or "to which we attend," *ad-tendere*) and yet could not do so, if not because of a mysterious participation already given to us in the ontological intimacy of mind and being.

This is why intelligence can never be "artificial": artificial intelligence is a preordained function wired to carry out certain operations. If we speak of intelligence, and can do so only analogically, we should always distinguish it from the intelligence of a person, a human being. Whereas a machine possesses, at best, "syntactical" capacity, the

⁴⁷ Emmanuel Levinas, *Of God Who Comes to Mind*, trans. Bettina Bergo (Stanford, CA: Stanford University Press, 1998), 146.

⁴⁸ Emmanuel Levinas, *Otherwise Than Being or Beyond Essence*, trans. Alphonso Lingis (Pittsburgh: Duquesne University Press, 1998), 148–49. "Inspiration" is existence "through the other and for the other, but without being alienation," (114–15).

person is capable of “semantic” appropriation—i.e., is able to understand what she is doing:

In the sense in which people “process information” when they reflect, say, on problems in arithmetic or when they read and answer questions about stories, the programmed computer does not do “information processing.” Rather, what it does is manipulate formal symbols. The fact that the programmer and the interpreter of the computer output use the symbols to stand for objects in the world is totally beyond the scope of the computer. *The computer, to repeat, has a syntax but no semantics.* Thus, if you type into the computer “2 plus 2 equals?” it will type out “4.” But it has no idea that “4” means 4 or that it means anything else. And the point is not that it lacks some second-order information about the interpretation of its first-order symbols, but rather that its first-orders don’t have any interpretations as far as the computer is concerned. All the computer has is more symbols.⁴⁹

Like Searle, Hubert Dreyfus has been especially critical of the artificial intelligence model of human thinking and cognition understood as disembodied processes. Overall, his is a critique of disembodied artificial intelligence. Dreyfus uncovers a number of false presuppositions entailed by disembodied artificial intelligence. First, a biological assumption: the brain processes information in discrete operations by way of some biological equivalent of on/off switches. Second, a psychological assumption: the mind can be viewed as a device operating on bits of information according to formal rules. Third, an epistemological assumption: all knowledge can be formalized; what can be understood can be expressed in terms of logical relations. Fourth, an ontological assumption: since all information fed into digital computers must be in bits, the computer model of the mind presupposes that all relevant information about the world, everything essential to the production of intelligent behavior, must in principle be analyzable as a set of situation free determinate elements. His conclusion:

Thus the view that the brain as a general purpose symbol manipulating device operates like a digital computer is an empirical hypothesis which has had its day. No arguments as to the possibility of artificial intelligence can be drawn from current empirical evidence concerning

⁴⁹ John R. Searle, “Minds, Brains, and Programs,” in Boden, *The Philosophy of Artificial Intelligence*, 85 (emphasis mine). Searle’s conclusion: “The point is that the brain’s causal capacity to produce intentionality cannot consist in its instantiating a computer program, since for any program you like it is possible for something to instantiate that program and still not have any mental state. Whatever it is that the brain does to produce intentionality, it cannot consist in instantiating a program since no program, by itself, is sufficient for intentionality” (“Minds, Brains, and Programs,” 87).

the brain. In fact, the difference between the “strongly interactive” nature of brain organization and the non-interactive character of machine organization suggests that insofar as arguments from biology are relevant, the evidence is against the possibility of using digital computers to produce intelligence.⁵⁰

Intelligence can only be an embodied act (think of the notion of “emotional intelligence”!), the actualization of knowledge and thinking through which we complete the world, we make the world come to itself, while receiving from the world the gift of a deeper sense of ourselves.⁵¹

The distinction between human and artificial intelligence becomes even clearer when we analyze basic dimensions of intentionality. Consider desire, imagination, and memory. In the eruption of *desire*, something possible only to an incarnate singularity, there emerges for the self the possibility to be other to itself. Desire reveals the energy of transcendence at work within the self and testifies to the power of self-differentiation in the self. There is more to the self than the identity of the same.⁵² When the latter feels itself as lacking, it not only expresses something negative but, rather, more positively brings forth the energy of being in all its richness, exploding in the self in the form of desire. As Levinas says: “Desire is an aspiration that the Desirable animates; it originates from its ‘object’; it is revelation—whereas need is a void of the Soul; it proceeds from the Subject.”⁵³

In this process, desire and *imagination* are allied. Imagination brings to further clarity the process of othering in the self, for desire acquires specific contours only in imagination. One ought to remember that the process of othering takes place in the intermediation with the other. The self is not fixed, it is a *metaphor*, a carrying across

⁵⁰ Dreyfus, *What Computers Still Can't Do*, 156.

⁵¹ On this, consult the beautiful essay by Herman Krings, *Meditation des Denkens* (München: Kösel, 1956).

⁵² For Paul Ricoeur to be a self is not to be the same, *idem* has to become an *ipse*! The issue of the relation between Ricoeur and Levinas is complicated; the two positions ought to be carefully distinguished. To put it briefly, and with reference to a pithy quotation from Ricoeur, the point of disagreement between the two consists in the fact that, for Ricoeur, “Awakening a responsible response to the other’s call cannot work except by presupposing a capacity for reception, of discrimination, and of recognition” (*Oneself as Another*, trans. Kathleen Blamey [Chicago: The University of Chicago Press, 1992], 339). See Cohen’s comment: “Following an existentialized version of philosophy’s transcendental route, for Ricoeur, in contrast, there must always first be self-reflexivity, a capacity in the sense of a base, ground, zero-point, from which and out of which and into which otherness is *correlated*. For Levinas, in contrast, such an insistence on recognition, or on recognizing the priority of recognition, misses accounting for the prior impact which is at once the impact of alterity as such and moral obligation” (Cohen, *Ethics, Exegesis and Philosophy*, 304–05).

⁵³ Levinas, *Totality and Infinity*, 62.

which differentiates itself in the images of itself it images for itself. The self is the metaphor of a carnal mindfulness.

Careful vigilance is required, though, since the process of imagining itself as other to itself is equivocal. It can be seen as an end in itself, a process in which the self goes from one self-formation to the next, never encountering the other, and in so doing, never coming to itself. The self is then in flight from itself and not toward the other.

We end up dissipating the original energy of being, given to us in who we are, when we cannot face the selves we are in promise. *We need to remember* who we are in our self-transcendence. *Memory* is needed to balance desire and imagination, as well as the quest for self-transcendence: "Memory recaptures and reverses and suspends what is already accomplished in birth—in nature...By memory I ground myself after the event, retroactively: I assume today what in the absolute past of the origin had no subject to receive it and had therefore the weight of a fatality."⁵⁴

Memory is the persistence of elemental self-awareness in the passage of transcending or becoming: "Memory as an inversion of historical time is the essence of interiority."⁵⁵ It is the return of the self to itself in the passage of becoming. I am talking here about memory in a non-objective way: not so much in terms of the process of remembering things, but as a kind of non-objective function, grounding our sense of interiority. Self-transcendence is not only externally directed: memory opens up the self to its inner otherness.⁵⁶ How is one to speak of artificial intelligence in terms of desire, imagination, and memory?

SECOND ETHICAL SELVING: ACTION

In action we come to ourselves as moral agents. This is first of all an openness, a response (*Wertantwort*) to the world and call of values which, for Levinas, is being revealed in the face of the Other, its vulnerability and indigence. To act morally is to transcend oneself; better, to embark in a movement of *transascendence*: "The metaphysical movement is transcendent, and transcendence, like desire and inadequation, is necessarily a transascendence."⁵⁷

This in two ways: we transcend ourselves in what we become, when acting morally. This is why the "doing" involved in acting (*praxis*, *agere*) is different from the doing of production (*poiesis*, *facere*). In the latter, we do something that brings forth an external being, an external object. In the former, the effect is not outside the

⁵⁴ Levinas, *Totality and Infinity*, 56.

⁵⁵ Levinas, *Totality and Infinity*, 56.

⁵⁶ I am thinking of the discovery of the unconscious or Dostoevsky's "underground man" in *Notes from the Underground*: the ground of autonomy turns out to be a groundless abyss!

⁵⁷ Levinas, *Totality and Infinity*, 35. The term "transascendence," as Levinas clarifies in the footnote, is from Jean Wahl.

agent, but on the agent itself. The effect of moral action is the achieved integrity of agency.⁵⁸ The robot produces effects, but does not act in the sense above, thus it cannot be a moral agent.

The transcendence in question, however, is ultimately toward the other as other. For this reason, there can be an ambiguity at play here, when the process of moral performance turns into a journey of self-achievement, whereby the other is sought only as a function of one's fulfillment. We can desire the other out of a lack, now seen not so much as impelled by the energy of the source that originated us, but driven by the sense of lack that, negatively, determines the other for-self.

One might call this kind of transcendence "self-oriented." Hegel understands the subject in the process of its becoming thus, as self-determining negativity: the other is for-self, like in the master-slave dialectic of the *Phenomenology*, a story that repeats itself in Sartre's dialectic of masochism and sadism.⁵⁹ Deformation takes place with the stifling of the plenitude of excess, which is the origin prior to the lack of the self-oriented self, culminating with Nietzsche in the will to power willing itself for the purposes of its own self-glorification. Perhaps, with Levinas, one can envisage another way. For him "metaphysics does not coincide with negativity."⁶⁰

THE PERSON AS OPENNESS TO THE OTHER

We need to understand the fulfillment of the self differently. While the self for sure is *penia* (poverty), it is also excess, because *porous* to the sourcing power.⁶¹ The openness of transcendence can be *agapeic* (other-oriented), rather than self-oriented: it can be an openness for the sake of other-being, rather than self-being. The "agapeic self" breaks the circle of mediation and returns to the origin (God?) as a source of infinite energy. Impelled by the generosity of the origin, the self opens itself up to what is other to itself, breaching the circle of self-mediation. In self-oriented transcending, the self is more than the transcendence. It is *transcendence*: in other-oriented transcending, transcendence is more than the self.

A dialectical mediation is at play, one that becomes an inter-mediation, rather than a sublation of the other to the self (*à la* Hegel). The space between self and other rests intact as the middle space between

⁵⁸ The point, with the distinction between *poiesis* and *praxis*, is obviously Aristotelean.

⁵⁹ See G. W. F. Hegel, "Independence and Dependence of Self-Consciousness. Lordship and Bondage," in *The Phenomenology of Mind*, trans. J. B. Baillie (New York: Harper & Row, 1967), 228–40; Jean Paul Sartre, *Being and Nothingness: An Essay in Phenomenological Ontology*, trans. Sarah Richmond (London: Routledge, 2018), chap. 3: "Concrete Relations with the Other."

⁶⁰ Levinas, *Totality and Infinity*, 41.

⁶¹ Think of the story of love in Plato's *Symposium*.

infinitudes: to the inward infinitude of the self, there corresponds the infinitude of the other. The mediation between such a plurality of infinities cannot be a self-mediation, in which one subordinates the other to itself. As the singularization of communicative being, the self *as self* cannot be reduced to will to power but will have to be understood as the *willingness to give itself* up for the other. Herein lies the paradox: loss of self is finding oneself. The “agapeic self” is dis-interested, in the sense of transcending self-interest into the middle (*interesse*). It is also *hetero-archic*: subject to the other, and because of this, a subject. In free obedience to the other, the self finally finds itself. For Levinas this openness to the other is the ultimate meaning of fecundity: “The I springs forth without returning, finds itself the self of an other: its pleasure, its pain is pleasure over the pleasure of the other or over his pain—though not through sympathy or compassion. Its future does not fall back upon the past it ought to renew; it remains an absolute future by virtue of this subjectivity which consists not in bearing representations or powers but in transcending absolutely in fecundity.”⁶²

The openness to the world, chiefly the world of the other, is an act of love, a fulfillment of reciprocity. To be a person is to love, because this is what “the incarnate singularity of a self, open to the Other” ultimately does: it actualizes itself beyond itself, in the responsibility for the other that is both a response (responsibility comes from *respondere*) and a release. A responsibility: I become myself, I become a moral agent, because I see myself commissioned by a call. “I am summoned as someone irreplaceable. I exist through the other and for the other, but without this being alienation.”⁶³

I can be many things, and yet fail myself, when failing to heed the particular call to which life calls me (this forms the nucleus of truth in situation ethics). Such call is singular: it is not a general call for “the humanity in me,” as Kant would have it; nor is it a response reducible to the production of a good, not even “the greatest good for the greatest number” of utilitarianism: “Freedom is borne by the responsibility it could not shoulder, an elevation and inspiration without complacency. The for-the-other characteristic of the subject can be interpreted neither as a guilt complex (which presupposes an *initial* freedom), nor as a natural benevolence or divine ‘instinct, nor as some love or some tendency to sacrifice.’”⁶⁴

Whence the moral call then? Why should one heed it? Why be moral, in the end? Kant put the question in terms of a difference between a “hypothetical” and a “categorical” imperative. He had a premonition of the issue at stake here but was ambivalent about recognizing

⁶² Levinas, *Totality and Infinity*, 271.

⁶³ Levinas, *Otherwise than Being or Beyond Essence*, 114.

⁶⁴ Levinas, *Otherwise than Being or Beyond Essence*, 124.

an alterity that summons the autonomy of the self, lest falling into heteronomy again. A communion of autonomous beings will be possible, for Kant, only in the "kingdom of ends." But this is only a postulate, an exigence of our thinking, not a reality we can know.⁶⁵

The question here is the question of God as the ground of morality. An important question, and not only for Kant.⁶⁶ At stake is not only the question of God as the law-giver, who grants the moral imperative its absoluteness. It is rather a question of *release*: the release of our own freedom into the reciprocity of love, out of the love that generates us into being. We love because we are being loved. Generated into love, in the space of goodness predisposed for our enjoyment, we are capable of a freedom-for-the-other beyond autonomy, in the generosity of service. This is to be a moral being. This is, ultimately, what it means to be fully human, to be a person.

If I see correctly, Levinas has a tendency to moralize the relation to God and, subsequently, the notion of creation. He sees God as the infinite *in* the face of the other, but *not as the ground* of love for the same, which opens the same to the other. Because the other is the Master, "The interlocutor is not a Thou, he is a You: he reveals himself in his lordship. Thus, exteriority coincides with a mastery. My freedom is thus challenged by a Master who can invest it."⁶⁷

There is moral earnestness in this God, but this is hardly a God of love! A different notion of creation is also needed. Levinas sees, correctly, that creation is a freeing of the person into her autonomy but does not quite understand such freeing against an ontology, a metaphysics of goodness. If this is the conclusion, then I am already *beyond* Levinas.⁶⁸

⁶⁵ Immanuel Kant, "Transition from Popular Moral Philosophy to a Metaphysics of Morals," in *Grounding for the Metaphysics of Morals*, trans. James W. Ellington (Indianapolis: Hackett, 1981), 43.

⁶⁶ Recall Ivan's argument in Dostoevsky's *The Brothers Karamazov*: "If you were to destroy the belief in immortality, not only love but every living force that maintain the life of the world would at once be dried up.... For every individual...who does not believe in God or immortality, the moral law of nature must immediately be changed into the exact contrary of the former religious law." Fyodor Dostoevsky, *The Brothers Karamazov*, trans. Constance Garnett (New York: Barnes & Noble Classics, 2004), 71. The quotation is in Book 2, chapter 6 ["Why Is Such a Man Alive?"].

⁶⁷ Levinas, *Totality and Infinity*, 101.

⁶⁸ As William Desmond puts it, "Levinas seems to reiterate again and again, not only the horror of the *il y a*, but the *evil of being* in the relentless self-insistence of the *conatus essendi*. ... His version of Plato's Good beyond being dictates a saving trauma and reversal from myself to the other. But *is there not an evil in that ethical good that sees being as evil?* ... We have not quite been released to the agapeic 'It is good'" (*Art, Origins, Otherness: Between Philosophy and Art* [Albany: SUNY Press, 2003], 162). I have articulated the implications of such a metaphysics of the good for the field of bioethics in Roberto Dell'Oro, "On the Ultimate That Is the First: Thinking Beyond (Bio) ethics," *Gregorianum* 100, no. 3 (2019): 621–47.

CONCLUDING UNSCIENTIFIC POSTSCRIPT

At the end of this long detour on Levinas and the philosophy of the person, I return to the initial provocation and the question “can a robot be a person?” What to make, now, of the dream of an immaterial universe, the mad project instigated by a notion of robotics in which the human being becomes, in the end, “the subjugated subject”?⁶⁹

In his 1920 novel *R.U.R.*, Czech writer Karel Capek describes the world imagined by scientist Rossum, a world of artificial workers (the word “robot” derives from the Slavic root for “work”) intelligent and indefatigable, but incapable of feelings: “Rossum’s Universal Robots,” hence the acronym for the novel’s title, “R.U.R.” Although the intention of Rossum is to liberate human beings from the slavery of work and make them the masters of creation, the dream eventually fails: once all their needs are satisfied, human beings no longer have to work, and thus cease to reproduce. The robots, however, revolt and kill the humans, and their leader proclaims, “The time of man is over. A new world begins. The reign of robots.”⁷⁰

We are not there yet of course. And there is no need to think that robotics as a scientific enterprise has to necessarily end in transhumanist madness. Still, the retrieval of a personalist philosophy capable of highlighting the essential difference between person and machine provides an important buffer to any scientific totalizing pretense. As long as this is the case, “the time of man is *not* over yet.”

Levinas’s phenomenological account provides important insights that remain closed to any scientific gaze, including the one articulated by Ishiguro and his vision of a “human-robot symbiotic society.”⁷¹ For Levinas, as for Merleau-Ponty, phenomenology “is from the start a forswearing of science.”⁷² Thus the “unscientific” nature of this conclusion, which only calls for the suspension of those premises that make it impossible to see the person’s singularity in a spectacle of world-objects. In this paper, I have attempted to provide something like an archeological reconstruction of such premises, only to show that the dualism they subtend falls short of accounting for what the person is: “an incarnate singularity, coming to itself, in openness to the Other.” M

Roberto Dell’Oro is the Austin and Ann O’Malley Chair in Bioethics, the Director of the Bioethics Institute at Loyola Marymount University, and professor in the Department of Theological Studies. He earned a doctorate in

⁶⁹ Thus Rémi Brague, in his reconstruction of modernity. See his *The Kingdom of Man: Genesis and Failure of the Modern Project*, trans. Paul Seaton (Notre Dame: University of Notre Dame Press, 2018), 160–68.

⁷⁰ See reference to Capek’s novel in Brague, *The Kingdom of Man*, 167.

⁷¹ Ishiguro, “Studies on Interactive Humanoids.”

⁷² Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Colin Smith (Oxford: Routledge, 1962), ix.

moral theology at the Pontifical Gregorian University in Rome and specialized in bioethics at Georgetown University. Roberto is the author/co-author of four books: *Pope Francis on the Joy of Love: Pastoral Reflections on Amoris Laetitia* (Mahwah, NJ: Paulist, 2018); *Health and Human Flourishing: Religion, Moral Anthropology, and Medicine* (Washington, DC: Georgetown University Press, 2006); *Esperienza morale e persona* (Rome: Gregorian University Press, 1996); and *History of Bioethics: International Perspectives* (San Francisco: International Scholars, 1996). He has translated Klaus Demmer, *Shaping the Moral Life* (Washington, DC: Georgetown University Press, 2000).