



## Ethics and cybernetics: Levinasian reflections

Richard A. Cohen

*Department of Religious Studies, University of North Carolina at Charlotte, NC, USA*  
E-mail: richachoen@aol.com

**Abstract.** Is cybernetics good, bad, or indifferent? Sherry Turkle enlists deconstructive theory to celebrate the computer age as the embodiment of “difference.” No longer just a theory, one can now live a “virtual” life. Within a differential but ontologically detached field of signifiers, one can construct and reconstruct egos and environments from the bottom up and endlessly. Lucas Introna, in contrast, enlists the ethical philosophy of Emmanuel Levinas to condemn the same computer age for increasing the distance between flesh and blood people. Mediating the face-to-face relation between real people, allowing and encouraging communication at a distance, information technology would alienate individuals from the social immediacy productive of moral obligations and responsibilities. In this paper I argue against both of these positions, and for similar reasons. Turkle’s celebration and Introna’s condemnation of information technology both depend, so I will argue, on the same mistaken meta-interpretation of it. Like Introna, however, but to achieve a different end, I will enlist Levinas’s ethical philosophy to make this case.

**Key words:** ethics, information technology, Levinas, neutrality

Meta-theorists of cybernetics<sup>1</sup> are prone to make large claims regarding the impact of this contemporary technological development. This is true of both Turkle<sup>2</sup> and Introna.<sup>3</sup> Their evaluations suggest that cyber-

<sup>1</sup> Richard Cohen uses the word ‘cybernetics’ to refer to all forms of information and communication technology – editors.

<sup>2</sup> See, Sherry Turkle, *Life on the Screen* (New York: Simon and Schuster, 1995). Turkle’s basic position, to express it in a preliminary fashion, is that the virtual realities made possible and apparently encouraged by technical capacities manifest on the computer screen, concretely illustrate and instantiate (a) a radical aesthetic relativism (or “freedom”) and (b) concomitant de-centered and multiple selves, constructed selves. What I have just called a “radical aesthetic relativism” is the ultimate “reality” theoretically defended by “post-moderns” such as Nietzsche, de Mann, Derrida, and Deleuze. What Turkle celebrates is what she takes to be the actualization of post-modern theoretical claims by means of the possibilities opened up for the first time by information technology. I would say that Turkle is no doubt right to see that they do illustrate the peculiar theoretical claims made by the post-modernists. But at the same time, contrary to Turkle, illustration does not constitute proof. In fact, so it seems to me, computer generated improvisations prove precisely the reverse of the post-modern claim: not that ultimate reality is relativist, multiple, fragmented, Heraclitean, as it were, but rather that the complex of these adjectives apply best to the limited field of aesthetic play now possible on computer screens. In the virtual reality of a whole slew of video games, for instance, a player has a virtual figure (boxer, martial artist, soldier, terrorist, et al) kill or pummel or otherwise violate dozens of virtual opponents every minute, but – thank God – not in reality.

<sup>3</sup> Lucas D. Introna teaches in the Department of Informa-

netics cannot simply be viewed from the perspective of morality, but rather that morality must be viewed from the point of view of information technology. Thus Turkle celebrates cybernetics’s ability to support a new form of selfhood, the de-centered and multiple self (or rather selves). The multiple self cannot be held accountable in the same way that the integral self of morality is held accountable. Cybernetics, then, liberates the traditional self for the freedom of multiple selves.

Introna, for his part, seems to be arguing the reverse point when he condemns information technology. But in fact he too credits cybernetics with a radical transformation, or the possibility of a radical transformation of morality. Because it mediates the face-to-face relation that, according to Levinas’s ethical philosophy, is the very source of morality, cybernetics would be the destruction of morality. So what Turkle celebrates Introna condemns, but both adhere to the same meta-interpretation of cybernetics whereby it is considered capable of radically transforming the human condition.

tion Sciences at the London School of Economics and Political Science. His position is partly articulated in Lucas D. Introna, Proximity and Simulacra: On the possibility of ethics in an electronically mediated world, *Philosophy in the Contemporary World* (forthcoming). My own knowledge of his position comes via our email correspondence. He was the organizer of a conference of a group called Computer Ethics: A Philosophical Exploration (CEPE), held at the LSE, on the topic “Ethics and Cybernetics,” on December 14–15, 1998, where I presented the first version of the present paper.

The question, as I shall pose it, then, is whether computer technology produces a radical transformation of humanity, or whether, in contrast, it is simply a very advanced instrument, tool or means of information and image processing and communication that is in itself morally neutral. If it is the latter, then far from freeing humanity for a new selfhood, as Turkle thinks, or ruining the moral foundation of humanity, as Introna thinks, cybernetics would, like all instruments, offer itself to both good and evil use. It is this latter view for which I argue in this paper. One consequence of taking the instrumental view is that the changes introduced by cybernetics – the so-called “computer revolution” – are not at all as radical, important or transformative as many of its proponents, such as Turkle, or its detractors, such as Introna, would have it.

First, a word of caution. Academic circles in the humanities tend naturally to be dominated by a humanist intellectual reflection and criticism. In our day they also tend to be dominated, at least in their rhetoric, by “leftist” views. Thus in humanistic academic circles one usually hears that information technology is not a great good but rather that it is intrinsically evil, somehow demonic, the arm of Big Brother. In this paper I will be opposing the view that cybernetics is either intrinsically good or intrinsically evil. But given academic prejudices, I caution readers not to confuse my instrumental position with a genuinely conservative or “rightist” view. My argument is that cybernetics, like all other human inventions, and much else, can be used for good or for evil – and is in fact already being used for both. Information technology is a tool. Just as the Bible can be used to support unjust wars, and guns can be used to save lives, so, too, cybernetics can be used and abused. This is another way of saying that computer technology, for all that it can do, has not effected any fundamental revolution in human affairs.

It seems to me that the academic “leftist,” “green” or “politically correct” view is one of the contemporary avatars of an unverified and unverifiable *gnostic* assumption that believes the material world, and most especially human interventions in that world, is evil as such. It seems to me, too, looking to the other side, that the “rightist” view, which takes cybernetics to be an unmitigated good, is naïve, optimistic and deluded by wishful thinking (if not motivated by economic self-interest).

To argue my case, I am going to bring to bear the ethical philosophy of Emmanuel Levinas. Admittedly, though he developed his philosophy over the second half of the twentieth century, Levinas did not specifically write about cybernetics, computers and information technology. (Neither did other great moral philosophers like Plato, Aristotle and Kant, a fact

that presumably does not render their theses irrelevant.) Nonetheless, it seems to me that his ethical metaphysics, which derives the “ought” of morality from the imperative force of the other on the I in the face-to-face relation, has clear and adequate conceptual resources with which to raise and deal with the question of cybernetics and ethics. Indeed, I think that Levinas’s ethics, precisely because it avoids the epistemological biases of previous ethical theories, is ideally suited to raise and resolve the question of the ethical status of information technology.

I introduce Levinas with what I take to be an especially pertinent citation. It appears at the center of his second great work, *Otherwise than Being or Beyond Essence* (1974), in a chapter entitled “Substitution.” Let us pause a moment first, however, to understand this term, substitution. Levinas thinks of moral selfhood not in terms of an agent’s freedom or autonomy, but rather in terms of the extreme proximity of the other. The moral self arises subject to the other. Indeed, such a self substitutes the other’s concerns for its own, to the point that one person can even die for another. The proximity of substitution, then, putting oneself in someone else’s shoes, making the needs of the other one’s own first priority, is the simple but profound gesture at the cornerstone of all morality. Indeed, it is not a gesture at all, if by this term we think of the self as a substance and substitution as an attribute. Rather, the proximity of putting the other first, being for-the-other before oneself, represents the very constitution of moral selfhood. It is the self constituted as moral sensibility rather than as animal self-preservation and self-interest or mindful self-knowledge. In giving priority to this self, the moral self in proximity to the other, Levinas is thus rejecting the privilege given to knowing, and the self-presence demanded by knowledge, by almost all of Western philosophy.

*The citation.* At the start of the chapter of *Otherwise than Being or Beyond Essence* entitled “Substitution,” Levinas summarizes the originality and task of his own philosophical project in a way particularly relevant to our topic, “Ethics and Cybernetics”:

In starting with sensibility interpreted not as a knowing but as proximity, in seeking in language contact and sensibility, behind the circulation of information it becomes, we have endeavored to describe subjectivity as irreducible to consciousness and thematization. Proximity appears as the relationship with the other, who cannot be resolved into “images” or be exposed in a theme.<sup>4</sup>

<sup>4</sup> Emmanuel Levinas, *Otherwise than Being or Beyond*

Differentiating sensibility qua proximity from knowing, and locating language within the former, i.e., language as “contact and sensibility,” which Levinas often names “saying” (*dire*), rather than as “circulation of information” or “thematization,” which Levinas names the “said” (*dit*), Levinas understands moral selfhood to be *irreducible* to thematizing consciousness. If we think of information technology in terms of “images,” “circulation of information,” and that which can be “exposed in a theme,” i.e., the *said*, then clearly it is the wrong category with which to think ethics, which in the first instance has to do with *saying*.

Quite naturally, Turkle and Introna also think of cybernetics in terms of information, images and themes. That is to say, in agreement with Levinas, and to use his terminology, they think of cybernetics in terms of the *said*. The failure of Turkle, however, is not to recognize that the significance of the *said* derives from its relation to an irreducible and more primordial *saying*. Her error is in fact exactly what Introna fears: losing sight of the moral significance of *saying* by reducing it away and totalizing the significations (circulation of information, themes and images) of the *said*.

What Turkle celebrates is precisely the immanent circulation of information. She celebrates meaning as a differential play of signifiers, variables bound by nothing other than their position within a language, their relation to one another. Cybernetics, by permitting for the first time a nearly complete freedom with regard to that positioning, is therefore, for Turkle, closer to the genuine character of meaning within language. Computers increase the freedom of signs to inter-play – intra-play – as signifiers within their synchronic/diachronic differential networks of signifiers. They also link up networks to form one network, one web. Computers, by permitting their users to come closer to the ideal of a free putting of all signs into possible relation to all other signs, permit the actualization, as it were, of the true nature of language and meaning. If by “life” we mean “meaningful life,” then the title of Turkle’s book, “Life on the Screen,”<sup>5</sup> summarizes her entire philosophy and her celebration of information technology. Cybernetics, then, would be the actualization of the radically aesthetic relativism theoretically expressed by such “post-modern” thinkers as Nietzsche, Wittgenstein, de Mann, Derrida, Deleuze and Rorty.

But Turkle has gone too far. She has succumbed to the typical philosopher’s error of confusing theory

with reality, indeed, of reducing reality to theory. No doubt she is right to suggest that the virtual reality constructed and reconstructed on monitor screens does – within limits – illustrate the peculiar relativistic theoretical claims made by deconstructive theorists. The recent movie “The Matrix,” for instance, attempts also to illustrate a “reality” artificially constructed “all the way down.” But its logic breaks down both within the narrative of the movie (incongruity of constructed bodies and mortal bodies), and in relation to those who made the movie and those who view it (incongruity of two dimensions and three). The illustrations Turkle celebrates, e.g. Multi User Dungeons (MUDs) and Internet Relay Chatrooms, like analogies, are only partially successful. They are not proofs. In fact, it seems to me that the illustrations favored by Turkle to demonstrate the “reality” of constructed selves and their constructed environments, prove precisely the opposite of what she intends. They show not that “ultimate reality” (or “non-ultimate unreality”) is relativist, multiple, fragmented and provisional, but rather that the complex of these adjectives applies only within the limited field of aesthetic play on the monitor screen (as the movie “The Matrix” only succeeds on the movie screen, and even there only partially or illogically). In the virtual reality of a whole slew of video games, for example, a player directs a representational figure (boxer, martial artist, warrior, terrorist, et al.) to kill, pummel or otherwise violates dozens of virtual opponents every minute. But these “violent” representations are not done – or “ought” not to be done – in reality.

In fact, and this will be Levinas’s point, one distinguishes virtual reality (computer monitor screens, movie screens, video screens, dreams, etc.) from “real” reality, not because one is an appearance and the other is a reality, as if ontology/epistemology could secure such a distinction. Rather this distinction hinges on the difference between the moral freedom of a fantasy world and the moral constraints of the “real” world. Though it may seem like the old but simple refutation of philosophical idealism, whereby one kicks a table or chair and declares, “this is real,” or one wonders what a “transcendental ego” has for breakfast, there is a profound difference between the “self” (or selves) constructed on a screen and the living and breathing self (and society) that constructs such a virtual “self.”

What Levinas contributes to this discussion is the recognition that philosophy’s appearance/reality distinction occurs *within* and is itself limited by epistemology/ontology. His claim is that if one remains within the field of epistemology and ontology one will indeed end up with a bottomless Heraclitean relativism. Not because such relativism is the case, but because one has overlooked the true “ground” of the

*Essence*, transl. Alphonso Lingis (The Hague: Martinus Nijhoff Publishers, 1981), p. 100.

<sup>5</sup> See note one above.

real, i.e., morality. Levinas's claim is that the difference between appearance and reality must be understood as the difference between an epistemological realm (the said, circulation of information, themes, images) and the exigencies of morality (saying, proximity, for-the-other, substitution). Moral seriousness and not some complicated but always deferred epistemology is the ground of reality. The flesh and blood self is mortal, suffers, and hence the I is responsible for the other. I am – to be a self is – responsibility for the other. On this basis one distinguishes virtual reality from “real” reality – screens from screams.

What about Introna? Unlike Turkle, he intends to reject deconstruction's virtual “reality” and aims to follow Levinas into morality. Thus selfhood must be understood more deeply than as a free construction. The moral self is constituted in the passivity of responsiveness to the alterity of the other, in proximity, in a responsibility to and for the other's suffering. All men are mortal, but the other's mortality concerns a moral self prior to its own: one is one's brother's keeper to the point that one can die for the other. Deeper than being for-oneself is being for-the-other, whose “being” is an ethical rather than an ontological condition. “Good” is thus moral responsiveness to the other, while “evil” is refusal of the other.

Accepting – or, as we shall, apparently accepting – these grounds, Introna argues that the thrust of information technology runs counter to morality. It induces individuals to interact not directly but mediately, i.e., across a barrier of information circulation, themes and images. Cybernetics would thus represent a refusal of the alterity of the other. It would thus be evil. Another way of saying this is to see that Introna agrees with Turkle's characterization of what occurs in cyber space communication. Cybernetic communication would no longer be a saying but a said. Face-to-face proximity would be replaced or covered over by the mediated and ultimately egotistical “life on the screen.” However ironic, “de-centered” or multiple one conceives such a life, because it eclipses and occludes the alterity of the other person encountered face-to-face, it undermines morality. Ultimately such “communication” would really be a solitary play, a private theatre, a game of mirrors, egoist even if the ego is fractured into millions of pieces. So while Turkle is right that information technology opens the possibility of a virtual self in a virtual world, for Introna this is precisely the evil of it. While Introna concedes that information technology *as an “empirical instance,”* i.e., as my particular word processing program, say, or my email account, or my web access, is indeed a tool, *as a “phenomenon,”* i.e., a “way of revealing the world” as such, it closes off face-to-face encounter, and hence is evil.

There is a peculiar irony about Introna's argument.

In the name of Levinas it defends the perspective of Levinas's greatest protagonist. That is to say, in its form Introna has invoked Heidegger's argument regarding the “danger” of technology. For Heidegger technology is not this or that mechanical or electronic device. Rather it is a way of revealing the world as such. Technology is the very character of our epoch. And what is that character? It is precisely the eclipse of the revelation of being. Technology is the “greatest danger” because it precludes any further developments in the issue that is of the greatest importance, namely, the ongoing revelatory character of the *Seinsfrage*.<sup>6</sup> Technology is pernicious, then, precisely because it is too successful, too pervasive, too profound, precisely because as such it blocks out the ongoing revelation of being. But Introna invokes Levinas. Recognizing the superiority of Levinas's ethical metaphysics to Heidegger's “fundamental ontology,” Introna understands that what technology eclipses is not the revelation of being but rather the revelation of the other.

But can Introna so simply superimpose the form of Heidegger's thinking onto Levinas's ethics? I think not. There is good reason for the Levinas/Heidegger opposition. The core of Heidegger's thought is precisely to think the epochal character of being, the *Seinsfrage*. It is, to be precise, a “turn” (*Kehre*) that allows being to reveal itself, allows being to be. Thus being today reveals itself as technological, even if being – as the thinker (Heidegger) realizes – has greater resources than its technological expression. The core of Levinas's thought, in contrast, is the moral responsibility of one person to and for another. From the point of view of this ethical metaphysics, being does not have the last word. Instead of Heidegger's “ontological difference” between beings (ontic) and the being of beings (ontological), Levinas will distinguish, as we have seen, between the sincerity of saying, saying in the face of the other, moral responsibility, and the congealing of such saying in the said, the refusal of the other. The said does not congeal because of beings/manifestations or the epochal being/manifestation of beings, but rather because one person, for whatever reason, has become insensitive to another. At the bottom of Levinas's thought is not the epochal power of being but the moral authority of human integrity.

Introna has, so I think, imported an illegitimate because incompatible Heideggerian notion into Levinas's ethical metaphysics. It seems to me that only if one succumbs to Heidegger's “Germanic” mode of

<sup>6</sup> See, Martin Heidegger, *The Question Concerning Technology and Other Essays*, trans. William Lovitt (New York: Harper and Row, 1977).

thinking, as Santayana would have understood it,<sup>7</sup> can one think of technology – in this case cybernetics – as having something like a life of its own, and hence an evil destiny. Introna, along with Turkle, deconstruction and Heidegger, takes cybernetic technology to be something fundamental and qualitatively new. Unlike Turkle and deconstructionists who celebrates this difference, but like Heidegger, who opposes it, he takes it to represent a qualitatively new threat, even if in his perspective it is a threat not to being's ongoing revelation but to the Levinasian morality of face-to-face encounter.

My argument against Introna, then, is that while I certainly sympathize with his fears, I do not think that he has fully appreciated the significance of Levinas's distinction between the said (circulation of information, themes, images) and saying. The very meaningfulness of the said depends on saying. If cybernetics were indeed something "new under the sun," and somehow permitted an unprecedented totalization of the said, then Introna, along with Heidegger (but for different reasons), would be right. It seems to me, however, that the much celebrated and/or lamented advances in the field of cybernetics represent not a qualitative change in Western spirituality, but a quantitative development.

Nothing persuades me that cybernetics is anything other than the most recent dramatic development in the long history of communications technology. No doubt this history is also the history of the ever increased possibility of anonymous communication, but at no point is there a radical break from the signature of the author. By comparison, I think the changes resulting from Johann Gutenberg's printing press in the fifteenth century, to take one example of another dramatic development, were even more far reaching. The printing press, with its movable type and its nearly exact reproduction, represented a great step in the development of the mass communication of information and images. It too was not a radical break, since it built upon several prior developments, the most significant of which was probably the invention of the alphabet. Like the invention of computers, it opened up new possibilities.

Prior to the printing press, the dissemination of information depended on handwriting, on hand copying. We think of the handwritten signature as the most personal of all written sign communication. But handwriting, too, can become less personal and more anonymous through calligraphy. The printing press

introduced the possibility of a greater anonymity and the possibility of a vastly greater distribution and availability of human communications. But even here the anonymity is not complete. Based on no less advanced chemical and microscopic technology, print as well as paper can be traced to their sources, just as a bullet can be traced back to the gun from which it was fired. Building on the developments initiated by Gutenberg, modern computers provide a *faster* means of storing, manipulating, duplicating and communicating signs. The difference is again quantitative. And again, like handwritten and printed letters, electronic email communication can be traced back to its source.

Cybernetics thus represents a quantitative development: increases in the speed, complexity and anonymity of communications already inherent in the printing press, an increase in the distance – but not a radical break – from the immediacy, specificity, singularity and proximity of face-to-face encounters. But carriages, trains, cars and airplanes also represent quantitative developments in speed and ease of travel. This is what technological development is all about: quantitative developments in speed, replication, miniaturization, and the like. I see nothing about information technology that makes of it a radically novel or unique menace – or boon.

Here is not the place to review the history of indirect communication, from smoke signals, tattoos and stele to hieroglyphs, alphabets and electronic blips. Plato in the *Phaedrus* had already presented Socrates' dual complaint about writing: it would both dull the memory and, more importantly, in comparison to "the living word" it "is properly no more than an image" (176). In the same dialogue, as elsewhere, Socrates had already distinguished beauty, which can be seen, from wisdom, which has no visible image (250). Writing, or life on the screen, would be lifeless and dull in comparison to the potential wisdom of the living word, let us say the "face-to-face." Writing would have come from Egypt, tainted by the preeminent cult of death. But even judged from the point of view of the living word, writing, too, would have its virtues – as Plato above all has shown by writing his dialogues. Plato still "lives" in his dialogues. No doubt Gutenberg's printed pages were met with both complaint and celebration, just as information technology is met today. But all communication – everything "said" – is "indirect." What is immediate, more immediate than knowledge, is the moral impingement of the other on the self, "saying" – both mine and the other's. One does not need alphabets or monitor screens to lose sight of the face. One can lose sight of the ethical face in the very flesh and blood face that faces. The "force" of the face is moral not ontological, one can kill, but one *should* not.

<sup>7</sup> See, George Santayana, *The German Mind: A Philosophical Diagnosis* (New York: Thomas Y. Crowell Co., 1968); originally published in 1915 as *Egotism in German Philosophy*.

Very simply, one can objectify the other. One can “interpret” the other’s face, “reading” from it symptoms, superstructures, ideologies, and the like. The face can always become a mask, become the distanced object of a psycho-analysis, discovering masks behind masks . . . in a process that goes on endlessly (deferral, task, program, future, etc.) or terminates arbitrarily in myth (“libido,” “Oedipus,” “mirror,” “archetype,” “class,” “gender,” “right brain/left brain,” “reincarnation,” etc.). But – and this is the key “but” – everything said indirectly points to a saying, a face, traced also in an author’s signature, in handwriting, in printing, and in email. The source of all signification lies not in signs relating to signs, in the said, but more deeply, more seriously, more painfully, in the moral significance of the face that obligates.

Email, the web, number crunching, problem solving, word processing, 3-D imaging, remain “consciousness and thematization,” to return to Levinas’s words cited above, even if that consciousness and thematization can be recorded and stored in vast quantities and rapidly communicated around the globe. Putting aside evaluations and superlatives, for mathematicians the computer, with its vast information storage capacities, remains a calculating machine; for writers it is a typewriter and secretary; for researchers it provides access to a global (and often unreliable) encyclopedia; for engineers and architects it is a graphic designer; for lovers and friends it is a new vehicle for distant conversation, a new way to stay in contact, to stay “in touch” when apart. By itself, however, like an abacus, a typewriter, or a telephone, but also like the printing press, the computer remains an instrument, a tool, an artificial aid to intelligence, but not a human intelligence at all. Unlike the human, it is neither good nor evil, though it can serve both.

In other words, the “evil” of cybernetics, to the extent that it is evil, is not specific to cybernetics. Nor is it even specific to technology or to an epoch of being, as Heidegger thinks. Evil is not the occlusion of being, but the refusal of the alterity of the other. No doubt it is true that information technology, like all developments in technology, makes a great anonymity possible, and hence increases the means of refusal. But at the same time, so it seems to me, it also increases the means of contact, of genuine communication, of inter-humanity. It is a trite saying that information technology has made the world smaller, meaning that it has brought people and things closer to one another. But trite as this saying is, it remains true – for better and for worse. Heidegger’s error, one that I think seduces Intra, is what I earlier, following Santayana, called a “Germanic” move: the effort to treat substance as subject, the anthropomorphizing of being. One sees this maneuver at work in Hegel’s *Geist*, Schopen-

hauer’s *World Will*, Nietzsche’s *Will to Power*, and in Heidegger’s *Ereignis* as well, even though the latter is invoked specifically to avoid “subjectivization.” It is as if being itself were alive, and in the case of its technological manifestation, as if it were an evil genius allied against us. But this manner of thinking is a mystification, a reification, projection and demonization of a spirit that is human (if not only human).

More than two decades ago, in a landmark book entitled *What Computers Can’t Do*,<sup>8</sup> (Boston: MIT Press, 1972; revised edition, New York: Harper & Row, 1979; reissued with new introduction under the title *What Computers Still Can’t Do*. (Boston: MIT Press, 1992). Professor Hubert Dreyfus persuasively argued that the calculations performed by computers cannot be called human thinking not because those calculations are less precise or less rapid than human thinking. Indeed they are or are becoming far more precise and rapid than human thinking. Today, only a handful of chess grandmasters can defeat a computer programmed to play chess. Rather, Dreyfus argues, computers do not think because computers lack sentient bodies and hence lack the background of common sense, know how and culture that an adult sensibility integrates into human thinking. Human intelligence, in radical contrast to artificial intelligence, would be essentially related to human embodiment, and to the horizons of human sensibility.

Professor Dreyfus’s thesis, somewhat disconcerting, apparently, for certain meta-theoretical partisans of artificial intelligence, is a particular application of the broader position elaborated earlier by Michael Polanyi in his book entitled *Personal Knowledge*<sup>9</sup> (1962). Even more, however, it relies for its argument on the phenomenological studies of Maurice Merleau-Ponty. In his great works, *The Phenomenology of Perception*<sup>10</sup> (1945) and *The Visible and the Invisible*<sup>11</sup> (1964, posthumous), Merleau-Ponty had shown in great detail that and how intelligibility and sensibility, sense and sensibility, are inextricably bound to one another, and both bound to larger cultural significations. For Merleau-Ponty sensibility is not abandoned in human thinking, no matter how abstract, but is rather reflected and traced there. Of course Merleau-Ponty, like Professor Dreyfus, also draws from the earlier phenomenological studies of Martin

<sup>8</sup> Herbert Dreyfus, *What Computers Can’t Do*

<sup>9</sup> Michael Polanyi, *Personal Knowledge* (London: Routledge and Kegan Paul, 1962).

<sup>10</sup> Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Colin Smith (London: Routledge and Kegan Paul, 1962).

<sup>11</sup> Maurice Merleau-Ponty, *The Visible and the Invisible*, trans. Alphonso Lingis (Evanston: Northwestern University Press, 1968).

Heidegger in *Being and Time*<sup>12</sup> (1927), and even more heavily on those of the founder of modern phenomenology, Edmund Husserl, from his vast writings, both published and unpublished.

Nor should we be forgetful that prior to the findings of the twentieth century phenomenological school, the greatest credit for the central insight into the irreducible sensible character of intelligibility – as Merleau-Ponty also acknowledges<sup>13</sup> – must be given to Henri Bergson (1859–1941). Bergson, in all of his many writings, from his dissertation published in 1889, entitled (in English) *Time and Free Will*,<sup>14</sup> onwards, made the groundbreaking arguments – against long and established traditions of both empiricist and rationalist philosophies – for the same point, namely, the inextricable link between mind and matter, sense and sensibility.<sup>15</sup> Before Bergson, Immanuel Kant, especially his insight into the incongruity of spatial counterparts, for instance the absolute difference between left hand and right hand, or an object and its mirror image, had also glimpsed this same link: the irreducibility of space or sensibility to an abstract or disembodied intelligibility.<sup>16</sup> But this paper does not intend or need to provide an adequate historical survey of the discovery of the irreducible role of embodiment (though I think it fair and important to point to the groundbreaking and still relevant studies of Henri Bergson for special mention).

Although all these thinkers, from Bergson to Husserl to Heidegger to Merleau-Ponty to Polanyi to Dreyfus (and of course others I have not mentioned), recognize the irreducibility of sensibility within intelligibility, none of them, however, recognized the genuine significance of the bond joining sense and significance. This blindness is as true of Merleau-Ponty as it is true of Heidegger. Both these thinkers had vision enough to see that because sensibility and

intelligibility are linked – whether as “mood” (*Stimmung*) or as the “I can” of motility – then so too are *historical* and *cultural* significations a necessary part of intelligibility. But they did not see what Levinas saw: that the ultimate link between intelligibility and sensibility is the one-for-the other, the I suffering for the suffering of the other, of *moral* sensibility.

Computers do not think not because they are constructed of plastic, metal, silicon, or even biological materials, and hence, lacking full human embodiment, have deficient ontological and epistemological horizons for thinking. With modern medical developments, humans, too, can and do have metal and plastic body parts. And who knows, computers will probably have more and more biological components.<sup>17</sup> Rather, and this is Levinas’s insight, lacking human sensibility and the social horizons of human sensibility means that a computer lacks the *ethical* sensitivity that makes sensibility and embodiment human sensibility and human embodiment in the first place. Oddly enough, then, computers do not think – are not human – not because they lack human bodies, but because like stones and animals they lack morality. They are indeed embodied, but their embodiment, unlike human embodiment, is not *constituted* – or “elected” – by an ethical sensitivity. Computers, in a word, are by themselves incapable of putting themselves into one another’s shoes, incapable of inter-subject substitution, of the caring for one another which is at the core of ethics, and as such at the root of the very humanity of the human.

“Subjectivity” Levinas has written, is “irreducible to consciousness and thematization.” The self is irreducible to its intelligibility not simply because it is embodied, that is to say, has a body, is material. After all, biologists, physiologists, physicians and morticians, among others, treat the human body as an object, and are not at fault. Rather, subjectivity is irredu-

<sup>12</sup> Martin Heidegger, *Being and Time*, trans. John Macquarrie & Edward Robinson (New York: Harper and Row, 1962).

<sup>13</sup> See, especially, Maurice Merleau-Ponty, *In Praise of Philosophy*, trans. John Wild and James M. Edie (Evanston: Northwestern University Press, 1963), pp. 9–33.

<sup>14</sup> Henri Bergson, *Time and Free Will*, trans. F.L. Pogson (New York: Harper and Row, 1960). See also, by Bergson (English title, French publication date): *Matter and Memory* (1896), *Creative Evolution* (1907), *The Two Sources of Morality and Religion* (1932), and *The Creative Mind* (1934).

<sup>15</sup> See my essay, chapter one, “Philo, Spinoza, Bergson: The Rise of an Ecological Age,” in *The New Bergson*, ed. John Mullarkey (Manchester: Manchester University Press, 1999), pp. 18–31.

<sup>16</sup> For text references to Kant and further discussion of his idea of incongruent counterparts, see Jacques J. Rosenberg, *From the Unconscious to Ethics: The Genesis of Psychoethics* (New York: Peter Lang, 1998), pp. 29–32.

<sup>17</sup> In certain figures of fiction, such as the policeman-robot in the movie “Robocop,” or the character called Data in the television series “Star Trek: The Next Generation,” one is left to wonder whether one is dealing with a machine with organic parts or an organism with mechanical parts. To decide this question is important precisely because one has moral obligations and responsibilities first to organisms, indeed to human organisms, before one has moral obligations and responsibilities to machines that serve humans or other organisms. If forced to make a choice, one saves the pilot and not the jet – or one should. A captain does not “go down with the ship” unless he or she has too; the captain’s obligation is to make sure that everyone else is safe first before deserting the sinking ship. Note: to give priority to moral obligations and responsibilities to humans is not to deny the bearing of moral obligations and responsibilities toward the non-human, whether organic or inorganic. It is rather to locate the true source of moral obligations and responsibilities.

cible to consciousness and thematization because the most basic significance of human embodiment pertains neither to ontology, i.e., its being, nor to epistemology, i.e., its knowability, but rather to *morality and justice*. What constitutes the very humanity of the human, and hence the humanity of human intelligence, is the significance Levinas names “proximity”: ethical relationship with the other, “the relationship with the other, who cannot be resolved into ‘images’ or be exposed in a theme.”

Human sensibility transpires neither as being nor as knowledge but as the sincerity of an undergoing, the trauma of a vulnerability, the passivity of a suffering sensitive to the suffering of others. It is because the human self can be for-the-other before being for-itself that it can be a human self in the first place. To be a self is to be for-the-other. Hence Levinas responds to the question of Cain: “Am I my brother’s keeper?” by answering incisively that the I is precisely a human I insofar as it is its brother’s keeper, that the I is at bottom an answering for the other, a putting of oneself in the other’s shoes. In this way the self is a vigilance, an awakening to the moral necessity of redressing the suffering of each other and all others. Ethics is therefore not a gloss on or an attachment to an already constituted being; it is the very beginning and source of being. Being does not derive, as we have said, from the epistemological contrast between appearance and reality, or being and truth. Rather, the epistemological contrast between appearance and reality derives from the sincerity of ethics. Only beginning with moral sensitivity, with the immediacy of obligations and responsibilities, with the “real life” of moral imperatives beneath or behind the “virtual life” of the screen, beneath or behind the circulation of information, is one able to discover a reason, a justification, and a purpose to what would otherwise become a purely aesthetic and theatrical circulation of information and images.

Despite the remarkable complexity of their operational abilities, their ready and nearly instantaneous linkage, and the extraordinary rapidity of their information exchanges, computers are neither in proximity with one another nor in proximity with us, in the manner in which humans are in proximity with one another. Artificial intelligence is not human intelligence because it lacks moral sensitivity. Humans, who make and run computers, who calculate and circulate information by their means, are human not insofar as they make, calculate, and circulate information, but precisely insofar as they are morally sensitive to one another, able to be concerned for one another, able to give not merely information but aid. The value or worth of information does not depend, ultimately, on its accuracy, that is to say, on the value of accuracy in and of itself. Rather, the value of accuracy, like all

the values appropriate to information, such as clarity, coherence, accessibility and corrigibility, derive their ultimate value from their worth to a humanity constituted by and committed to the twin imperatives of eliminating evil and promulgating good, hence the imperative to create a just society.

The humanity of the human does not arise from an animal or a machine evidencing logic or the rationality of means and ends. Ants, termites, bees and porpoises, after all, are rational in this sense. Rather, the humanity of the human arises when an animal, or any being, is moved not by efficiency but by morality and justice. A being becomes moral and just when in its very sensibility, and across the pacific medium of language, it finds itself desiring an undesirable and insatiable service for the other, putting the other’s need before its own. The human is not a machine that is programmed with the words of the Bible or other sacred texts. Brotherhood, solidarity, the humanity of the human, is more concrete, more immediate, and more painful still: putting the other before the self. To be sure, humanity requires the order of logical and efficient intelligibility in order to create a just society, in order to create prosperity to feed the hungry and build homes for the poor. A moral humanity calculates and enlists machines to serve justice. If it happens that one day animals or machines become capable of independent moral sensitivity, then they too will enter into the unitary and unifying solidarity of moral agency. It is the priority of ethics that gives sense to the universe, indeed, that gives sense to the very notion of priority. This is Levinas’s point and this is the point of ethics. The greatest, noblest, most glorious cause, the most compelling imperative, is neither ontological nor epistemological, but ethical. Being and knowing *ought* to serve goodness, or they serve no purpose at all.

I have mentioned the possibility of animals and machines joining the one brotherhood of ethical sensitivity. In our day, however, moral responsibilities and obligations have their source in human sensitivities, in the humanity of the human. It begins in the human, to be sure, but from there it spreads to all creation, organic and inorganic. In our day, the ethical dimension of human proximity transpires across the communications made possible by computers, just as human proximity takes place across phone calls, letters, artifacts. The “face” can be a letter. The “face” can be an email message. The computers themselves, like alphabet letters and telephones, like pencils and books, however, are neither good nor evil. The “face” ruptures them, pierces them with the alterity of the other. By themselves they are shadows of shadows or masks of masks. For all the above reasons, the issues raised regarding “computer ethics” are at bottom the issues of ethics *simpliciter*.



Let there be no doubt that information technology can serve good. The global transparency it provides enables aid to be brought to natural disaster victims who might otherwise be unnoticed. It permits transmission of specialized medical procedures to remote hospitals. It permits transmission of vital and up-to-date meteorological data to farmers everywhere. It opens the doors of friendship and cultural exchange across political borders – to name only a few of the benefits of information technology. Let there be no doubt that computers can also serve evil: invasion of privacy; copyright theft; “anonymous” hate messages; cultural imperialism and leveling, and the like. But the computer and its computational abilities, information *technology*, is neither to be blamed nor praised. Humans relating to humans are good or evil, sensitive or insensitive, sacrificing or selfish. They are good or evil, or both, not because of this tool or that, however sophisticated the tool, but rather because the humanity of the human is constituted across a human sensitivity, a sensibility sensitive to and suffering for the suffering of others. We are free to respond to the other or to turn away, and if we choose to turn away we can no more hide behind our computer screens than behind Mercedes or manor doors.

Both the incredibly sophisticated manipulation of information and the nearly instantaneous global transparency provided by computers offers all of us a great opportunity for increased moral sensitivity, in each individual and around the world. It give us a new opportunity to become aware of and to share in the awesome and daunting obligations and responsibilities of a *global humanity*. And although I would like to think of increased transparency as an unmitigated good, I must admit, based on history and on the political experiences of the twentieth century most especially, that the same increased abilities to manipulate and disseminate information create powerful totalitarian opportunities too. But this is yet another way of saying that it is up to us, in the transcendence that is our proper humanity, to decide toward which direction, be it up or down, that we wish to use and steer the awesome power of contemporary information technology.

## References

- Henri Bergson. *Time and Free Will* (Translated by F.L. Pogson). Harper and Row, New York, 1960.
- Richard Cohen. Philo, Spinoza, Bergson: The Rise of an Ecological Age. In John Mullarkey, editor, *The New Bergson*, pp. 18–21. Manchester University Press, Manchester, 1999.
- Herbert Dreyfus. *What Computers Can't Do*. MIT Press, Boston, 1972; revised edition, New York: Harper & Row, 1979.
- Herbert Dreyfus. *What Computers Still Can't Do*. MIT Press, Boston, 1992.
- Martin Heidegger. *Being and Time* (Translated by John Macquarrie & Edward Robinson). Harper and Row, New York, 1962.
- Martin Heidegger. *The Question Concerning Technology and Other Essays* (Translated by William Lovitt). Harper and Row, New York, 1977.
- Lucas Introna. Proximity and Simulacra: On the Possibility of Ethics in an Electronically Mediated World. *Philosophy in the Contemporary World* (forthcoming).
- Emmanuel Levinas. *Otherwise than Being or Beyond Essence* (Translated by Alphonso Lingis). Martinus Nijhoff Publishers, The Hague, 1981.
- Maurice Merleau-Ponty. *Phenomenology of Perception* (Translated by Colin Smith). Routledge and Kegan Paul, London, 1962.
- Maurice Merleau-Ponty. *The Visible and the Invisible* (Translated by Alphonso Lingis). Northwestern University Press, Evanston, 1968.
- Maurice Merleau-Ponty. *In Praise of Philosophy* (Translated by John Wild and James M. Edie). Northwestern University Press, Evanston, 1963.
- Michael Polanyi. *Personal Knowledge*. Routledge and Kegan Paul, London, 1962.
- Jacques J. Rosenberg. *From the Unconscious to Ethics: The Genesis of Psychoethics*. Peter Lang, New York, 1998.
- George Santayana. *The German Mind: A Philosophical Diagnosis*. Thomas Y. Crowell Co., New York, 1968. Originally published in 1915 as *Egotism in German Philosophy*.
- Sherry Turkle. *Life on the Screen*. Simon and Schuster, New York, 1995.

