Article

Agency, Responsibility, Selves, and the Mechanical Mind

Fiorella Battaglia [1,2]

---

1   Faculty of Philosophy, Philosophy of Science and Religious Studies, Ludwig-Maximilians-University, 80539 Munich, Germany; Fiorella.Battaglia@lrz.uni-muenchen.de
2   Institute of Law, Politics, and Development, Sant'Anna School of Advanced Studies, 56127 Pisa, Italy; fiorella.battaglia@santannapisa.it

Abstract: Moral issues arise not only when neural technology directly influences and affects people's lives, but also when the impact of its interventions indirectly conceptualizes the mind in new, and unexpected ways. It is the case that theories of consciousness, theories of subjectivity, and third person perspective on the brain provide rival perspectives addressing the mind. Through a review of these three main approaches to the mind, and particularly as applied to an "extended mind", the paper identifies a major area of transformation in philosophy of action, which is understood in terms of additional epistemic devices—including a legal perspective of regulating the human–machine interaction and a personality theory of the symbiotic connection between human and machine. I argue this is a new area of concern within philosophy, which will be characterized in terms of self-objectification, which becomes "alienation" following Ernst Kapp's philosophy of technology. The paper argues that intervening in the brain can affect how we conceptualize the mind and modify its predicaments.

Keywords: brain–computer interface (BCI); human–robot interaction; mind; sense of agency; alienation

1. Introduction

The changes brought about by the use of neurotechnology informed by artificial intelligence (AI) can be dramatic, as Brain/Neural Computer Interfaces (BNCI) and robotics technologies can directly intervene in the brain and disrupt cognitive processes responsible for consciousness and identity raising questions about what it means to be a unique person. Our ordinary accounts of intentional behavior appeal to desires, habits, knowledge, reasons, and perceptions. These new interventions operate in a conceptual framework dominated by sub-personal categories. Such technologies are therefore a source for investigation in the field of moral theory. I propose that there are two possible paths of research in this area. On the one hand, questioning the impact of neurotechnology and AI includes philosophical, ethical, regulatory, legal, and social implications—an approach focusing on their applications. On the other hand, questioning neurotechnology and AI also includes concerns about how we make sense of ourselves and modify theoretical and practical concepts. By understanding and conceptualizing us in new, unexpected ways, we re-ontologize ourselves. The new philosophical framework cannot remain without practical implications as the new insights generated by this activity will trigger and motivate our actions. Therefore, as human enhancement technologies will make new claims, which impact competing approaches to the mind, they also represent a relevant area for moral theory. In this sense, neurotechnology motivates a relevant area of moral theory.

The philosophical framework behind this fascinating scenario was first named by James H. Moor in 1985 [1]; it is known as the anatomy of computer revolution, since he was using two stages for making explicit the perspective on the development of a cultural innovation initiated by the "introduction" and "permeation" of computers in social life. While the introduction stage does not alter social institutions and easily accommodate technology into the means-end relation, the permeation stage does have the potential to

disrupt social institutions and may bring unprecedented systematic thinking to bear on the very phenomenon of technology. His insights have undeniable interest and applications for theorizing about our merger with machines. It will be important for my purposes in this article to concentrate on the impact on moral theory. Therefore, I am not going to discuss in the paper the ethical issues related to the alterations following the interventions in the brain [2–5]. I will rather discuss some implications that these interventions might be able to trigger on scientific representing and understanding of the mind. The involved problems are not just ethical problems, but conceptual and philosophical ones.

While the first investigation will be concerned with challenges to human abilities, the latter investigation is concerned with the conceptual framework that these interventions are going to modify. My argument is focused on knowledge and aims to establish that interventions in the brain involve new understanding of the mind and on the basis of these new insights on the motivations for actions [6]. In other words, human enhancement technologies and our merger with machines call for the assessment of the adequacy of our representational devices for grasping what it means to be human from a moral, ethical, and legal standpoint. It rests on the idea that having access to the neural level of the brain might have an impact on both the knowledge about how it feels to have the experiences of that being and the knowledge about the fact that we evaluate each other's behavior and attitudes. It also challenges the idea that having physical knowledge about the mind might render redundant other sorts of knowledge such as the phenomenal and moral knowledge about mind and behavior. In a sense, inquiring features such as consciousness, intentionality, meaning, purpose, thought, and value; leads us to questioning whether they can be accommodated in a naturalistic image made of physical facts, whose access is granted by the physical sciences, this questioning renews a number of well-known philosophical disputes. Distilling this approach down even further, we might say that it aims to address the classical mind–body problem.

The mind–body problem has been addressed in a variety of ways. However, in the last decade, impressive advances have been made in the field of technological improvement of sensory, motor, and cognitive performances. This field, which has the potential to heal and enrich the mind, may also revolutionize the way in which the mind is understood and may have a bearing on how to model the mind. The reasons for avoiding a naturalistic approach to the mind are not reasons to avoid addressing the scientific treatment of it. A naturalistic approach to the mind, philosophically constrained by a compatibilist requirement (that free will is consistent with universal causal determinism) and supported by empirical evidence provided by the interplay with new technologies, generates a philosophically adequate and recognizable basis for a better understanding of the mind. This, in the most basic terms, is the thesis of this paper. Hence, the renewal of concerns against physicalist theories of consciousness and agency is currently being challenged by disruptive research and innovation, among which robotics and neurotechnology are prominent examples. At the metaphilosophical level, as technology is becoming an integral part of social practices, a relevant area of moral theory is consequently called into question. This is not a bad thing; on the contrary, it allows us to rethink how we know the characteristics of human mind. In this essay, I will present an analysis of the epistemic and ethical challenges posed by smart technologies about the way in which we understand the mind represented in its practical aspects as well [6,7]. Through a brief review of the three main approaches to the mind, first person perspective, second person perspective and third-person perspective, I will identify a main area of concern, deriving from accessing the mind and its features through neural mechanisms.

A survey of the literature in different fields provides a variety of suggestions: it is to be noted that the relevant impact is characterized in terms of enhancement technologies, new empirical evidence and transformation of vulnerability [8–13]. The nature of each of these categories, and the relationship between them, are both controversial. For example, some philosophers will question the improvement brought about by the enhancement technologies, while others will include thinking about human enhancement technologies as

a very broad tent, including their impact on our epistemic sources concerning our sensory, motor, and cognitive abilities. Finally, others will think of enhancement as a shift in our vulnerable features. For the sake of space, this essay will focus on the last two points throughout. This opens up new epistemic improvements for a re-appraisal of notions such as agency. At the same time accessing neural mechanisms to modulate human behavior makes individuals vulnerable to "nudges", manipulation, and deception, which will be made practicable, not least by the adjustment of our representative features.

The final aim of this paper is to draw a critical and constructive analysis and to show that by merging with a machine we have as a result both a surplus of representational devices and a vulnerability warning. I will first describe this surplus of representational devices in terms of direct knowledge of our actions, which has been fed by empirical studies. I will then address this kind of concern in term of "alienation" following Ernst Kapp's philosophy of technology [14]. The claim is that "alienation" should inform the discussion on transformative effects of these technologies, that is, on those effects which cannot always be interpreted to clear cases of epistemic or ethical shortcomings. Thus, this argument is conducive to a different kind of epistemic and ethical discussion resulting from the facts obtained by the ongoing technological transformation of human motor, sensory, and cognitive traits. It is the objective of this paper to clearly distinguish between the various thematic strands (consciousness, agency, mind), which are also dealt with in quite different debates—philosophy of mind, theory of action, theory of rationality and the theories of personal identity, theory of cognitive science and to clearly show the connection between them, which is not drawn by the single theories.

2. The Limits of Treating the Mind as an Object

Let us start with the way in which we represent the mind to assess what kinds of transformations have been occurring by our merging with the machines in our representational devices. How do we conceptualize agency, responsibility, and consciousness? We normally can count on three competing approaches: First, the approach articulated in first person perspective and tailored to fulfill the requirements both of subjectivity and of consciousness as experience [15–18] second, the approach articulated in second person perspective [19–21] and tailored to serve the needs of the personality; finally, the approach articulated in third person perspective and tailored to fit the human brain facts [22].

The distinctive features of the phenomenological mind and of the personality can best be seen by contrasting them with one more approach, the physical approach. The quality of what "it feels like" to be in such a state cannot, in principle, be conveyed in a physical language tailored to fit the objects in the world. Such facts of experience show an irritating incompleteness of the objectifying description of the world. The same happens with the personality layer and its normative requirements. The problem with the objectification of the mind is that it exhibits two kinds of shortcomings. It cannot account for consciousness as experience [23] and cannot account for the causal role of mental events in the physical world [24]. Therefore, the process of objectification—while proving fruitful in interpreting, explaining and predicting a series of physical events—shows its limits especially in the case when we want to accommodate features of personality along with those of subjectivity. When ascribing someone's agency or moral responsibility what matters is whether the person in question is able to participate in normative discourse and practice. Whether she is able to give and take reasons for her actions, beliefs, and feelings. On the one hand, we rely on these two layers of complexity—subjectivity and personality—when representing normative relations between agents and accounting for meaningful experiences of the world. On the other hand, the physical approach makes it possible to treat the mind as an object. As a result, it is possible to make various predictions. The underlying idea is that rooting the study of the mind in the natural sciences will provide us with a comprehensive understanding of the main mental features. Moreover, motor, sensor, and cognitive abilities seem to be accommodated in this naturalistic version of the world. This is the case at least

in psychophysical reductionism that says how the physical sciences could in principle provide an account of everything in the world.

The neural correlates of consciousness that cause the experience seem to be able to tell the whole story about the mind. A neural correlate of consciousness can be described as a minimal neural system that is directly related to states of consciousness. At least physicalist positions are keen to endorse this view. As fascinating as this kind of approach can be, as it promises to make the study of the mind scientific, it has costs, and I hope to show that they are not worth incurring. I will argue that there are philosophical costs to denying subjectivity and personality. I believe that physicalist approaches may allow for more philosophical insight, however. In this article, I will attempt to articulate what "getting at the mind" consists of, and what its merits and its faults are. A critical view says that psychophysical reductionism is a failure; the prospect that physical sciences could in principle provide a theory of everything is a failure in terms of explaining free actions and experience. In particular, the attempt to go from phenomenology to information processing in the brain is not going to happen in my view. Subjectivity and personality will be resistant to this kind of view. But interestingly, human enhancement technologies have introduced new insights in this discussion. By accessing the mind at the neural level, the human enhancement approach has introduced some major transformations in our knowledge.

There is a difference between being a mind and being an extended mind. It is an epistemic difference. The argument presented in this paper focuses on that difference. It is the distinct knowledge underling this specific self-objectification, which allows for detecting, influencing, and stimulating mental states. What presents itself in its relentlessness is the fact that the third approach may not be able to explain the other two dimensions elaborated in second- and first-person perspective and yet it does contribute to re-articulating a number of cognitive and affective features of the mind such as feeling, mood, perception, attention, understanding, memory, reasoning, and coordination of motor outputs and therefore high-level expressions of these properties such as selves and personality. In particular, the amount of knowledge available to be processed is remarkable if we consider the human–machine interaction. As a result, we are faced with a twofold novelty: some features of human behavior—prominently the "sense of agency", i.e., the kind of immediate knowledge we have of our actions– can be traced back to underlying mechanisms conceived as a causal chain of physical events at a sub-personal level and can be implemented and altered by accessing these very mechanisms at the neural level.

The debate about the "sense of agency" has been driven by empirical studies in the field of psychology and cognitive neuroscience [25–27]. What is more, robotics inspired by the idea of the extended mind has been another driver of this transformation. It says that the extended mind relies on sensorimotor synergies, and that only the understanding of their mechanisms will enable designing and building artificial systems that can generate a true symbiosis with the human. While we have learned much on both motor and sensory synergies in separation, not much attention has been paid before as to how the concept of synergies extends to the brain–body–environment interaction, and to the sensorimotor loop itself.

Pioneering work in the psychophysics and neurophysiology of the human hand and upper limb, and in the translation of this understanding in bionic concepts has been achieved [28]. The next level is to exploit the extended mind approach to a whole brain–body–environment interaction schema. In doing so, it will be possible both to explore the mechanisms by which the human and artificial system can remotely communicate and cognitively merge and to improve our comprehension of what it means to be human. This will in turn produce knowledge about our own actions. Whereas interacting with objects in the environment in which we live in is mostly immediate and unreflective, to engineer this very natural interaction (among our body, objects, and environment) is complicated. The success of engineered intuitive natural human–robot interaction (HRI) will critically depend on whether the operator will have no difficulty in recognizing the robot, its movements and sensations as part of their own body and their own actions: the

exploitation of systems such as these requires the fusion of the human mind with robotic sensors and actuators. If this is going to happen, by interacting with the robot, then the operator will be released from the physical, cognitive and affective workload associated with traditional teleoperation. Crucially, such a system relies on a multilevel approach to ensure engineered actions to be as successful as the natural ones. Not only will our skilled capacities be improved, but our understanding of natural and intuitive human–robot interaction makes room for the improvement of our knowledge about our own actions and sense of agency. This opens up three levels of inquiry, which fully cover the interaction with the robot and hence provide both explanations on how intuitive natural human–robot interaction occurs and instructions about how to ergonomically and ethically design the robot platform. The levels, in which we choose to describe, explore, and discuss the system and its context, are the neural level, the ergonomic level, and the ethical level.

The benefits of the proposed exploration of the extended mind idea and its robotics applications are manifold. It is common to think of psychophysical reductionism as a failure or as a promise. This way of thinking certainly has its merits, if only to a limited extent. For, unfortunately, the nature of the mind–body problem is not well-understood philosophically, despite the fact that some important past debates have discussed a remarkable number of issues. There still remains the advantage, however, that the concept of the extended mind is empirically amenable and therefore less rich, obscure and slippery than that of mind, and hence easier to handle. So, a definition of mind grounded in the symbiosis with technological artifacts seems to be a good starting point. The mechanisms of the mind, which basically rely on causal explanations, cannot make sense of subjectivity and personality. Intervening in the brain will modify their physical extension and will result in an impact both on the physical and phenomenological level. This is very much in line with the version of the identity theory contended by Davidson [24]. It says that all mental events are identical with physical events. However, not all physical events have a mental extension. We cannot make sense of mental events by recurring to physical science. This is often referred to as the nothing-but reflex [24]. This construct characterizes the aptitude of sentences like these "Conceiving the "Ode to Joy" was nothing but a complex neural event".

By relating mental events to other mental events, we explain them. This occurs by appealing to desires, intentions, and perceptions of the human agent. Explanations such as these represent a conceptual framework dominated by reasons and are usually detached from sub-personal level instantiated and explored by physical science. And yet, there are some particular cases. In cases where we know particular identities and are able to correlate particular mental events with particular physical events, we can modify the mental events by intervening in the brain. For example, Miller et al. [29] were able to find the dynamics of sensorimotor cortices associated with the touch of a hand-held tool. When the identity is not questionable, then we may be able to obtain a real advance. Epistemic and practical challenges like these call for a reappraisal of our epistemic setting. There are not only feasibility questions involved such as "how can we facilitate the symbiosis between machines and us?" As more difficult questions begin to have positive answers, healing and protecting the mind, understanding the mind, enriching the mind, and modeling the mind can be correlated with the way in which we make sense of our experience. After two decades of research, the neural level of the human-like interaction with artifacts starts to be disentangled. First advances are registered concerning motor abilities, and much of the functioning of the motor system occurs without awareness. Nevertheless, we are aware of some aspects of the current state of the system and we can prepare and make movements in the imagination. These mental representations of the actual and possible states of the system are based on two sources: sensory signals from skin and muscles, and the stream of motor commands that have been issued to the system. Human-like interaction with the robots can lead to the augmentation of the motor system in the awareness of action as well as improvement in the control of action. At the same time this very enhancement adds another piece to the "mechanical mind" as I call this new epistemic device [30]. At some

extent there is a convergence between "extended mind" and "mechanical mind". While the first is meant to connote the human–machine hybridization, the latter is supposed to mark the possible explanatory device arisen from this human machine hybridization. Blurring the line between persons and things will not be conducive to epistemic opacity. On the contrary, it will lead us to new insights.

Drawing on these new results, there are remarkable insights in the field of moral theory as well. It has been necessary to wait for the most recent developments of embodied cognition to give empirical evidence to the bodily involvement of the agent in the theory of action. The construct of the "sense of agency" provides new elements to overcome the vision of an agency characterized in the absence of bodily mechanisms of sensory processing and motor control. It provides a direct knowledge of our actions based neither on observation nor inference and can be implemented in a computational model of the agent. In addition, robotics fits well into this picture, because by investigating what happens at the neural level when we add an extension to the agent's body, it gives the possibility to modulate and correlate the knowledge related to the phenomenological experience with that related to the neural level. To combine a theoretical approach with a radical approach that relies on the construction of complete intelligent systems with a real apparatus of perception and movement that interacts with the environment is able to counteract misleading results. This approach presents strong evidence that the sense of agency originates in the neural processes responsible for the motor aspects of action. I contend that these recent results are compatible with a causal theory of action as suggested by our ordinary attitude that relates us to objects with a practical rather than simply theoretical stance. I believe that this contributes to a compatible perspective capable of integrating the knowledge that comes from different approaches to the same phenomenon. Giving scientific consideration to the physical events does not necessarily mean surrendering to naturalism; rather, it means not committing us in an alleged tension between two cultures—the scientific and the humanist.

3. Additional Epistemic Devices

In light of the fact that the traditional critical view has been revised and that accessing the neural correlates no longer means to subscribe to some form of reductionism, I will now argue that to integrate both the experiential dimension of consciousness and the space of reasons, in which we operate by appealing to reasons in order to explain actions, to make sense of our experience and to make room for normative claims, is fully compatible with the idea to treat our mind as an object. In short, the three rival accounts discussed above may not be competing against themselves anymore. The extra ingredients are the human enhancement technologies, which while allow our merging with the machine will produce an advancement in the understanding of the mind in terms of the same subject thought as necessarily united in its two components—the human and the machinic.

What is more, this kind of integration between human and machine will make room for new representational devices derived from our interaction with enhancement technologies. Its implied greater capabilities are conducive to more ethical sophistication in a conceptual framework. At this point in the discussion, I will now introduce some insights into human brain enhancements from a legal perspective. By looking to legal theory rather than moral theory, I may be able to better make sense of the transformative nature of human enhancements. As Brain/Neural Computer Interfaces and transparent robotics realize a seamless symbiosis with the human brain, they paradigmatically introduce new elements to be weighed in legal theory. A legal analysis will be able to elucidate some aspects of this new perspective, which I consider from a philosophical perspective in a second step of my analysis. Legal scholars seem to comply with the experiment to treat the mind as an object and also suggest that we need to re-shape the ontology accordingly. This is because, in the legal perspective, the main interest is a practical one. Mostly, it is a matter of regulating the human–machine interaction in the case that something goes wrong. This perspective, far from being departmental, turns out to be much more liberal, because it allows us to think

of personhood outside metaphysical burdens. The legal perspective, while aims to address how to deal with regulatory issues, develops a conceptual and regulatory framework able to overcome the dualism between humans and things.

The nature of this interaction is also a fascinating philosophical matter, and human–machine interaction from a metaphysical perspective raises interesting questions of its own. Fundamentally, they call for a re-examination of our understanding of the mind and its properties. It is necessary to make a detour to value and integrate these new elements into philosophical analysis. If this is not performed, the philosophical analysis remains stuck with two equally harmful prongs, scientific naturalism and antireductionism. The two major recent developments in philosophy of mind—naturalism and antireductionism—are deeply opposed to each other in important ways, but there is a striking convergence on answering the question—even though differently—whether scientific facts are able to properly account for consciousness, intentionality, meaning, purpose, thought, and value. As I said, there is another detour available with a view to obtaining a re-examination of familiar, perhaps too familiar, ways of thinking.

For Jens Kersten, there are no metaphysical standards that loom over the way we conceive of the mind [31]. Jens Kersten succeeds in combining elements of the theory of the nature of technology with elements of personality theory. In short, he is committed to the idea that technology as an element of culture may offer the possibility to reflect and identify with one's own artifacts. According to Jens Kersten, the symbiotic connection between human and machine is better understood as the very expression of personality. In his view, over our merger with the machine lies the decisive difference between the instrumental and the symbiotic constellation. The instrumental constellation is characterized by the differentiation between the human being as person and the machine as a thing. In those cases, in which the machine is essential for the development of personality, this relationship can be charged according to legal personality provisions, without, however, overturning the fundamental distinction between person and thing. In the symbiotic constellation, this differentiation between person and thing is omitted: The machines become "Nobjects" [32], which are going to get rid of their legal status as things and, due to the claim for dignity and the right of personality of their bearers, are going to be protected as part of their human personality. In this way, the prostheses, the implants, and other devices are going to acquire in a parasitic way the same moral status of their human host. As already stated, applicable legal theory is mainly concerned with a practical solution of those cases in which something goes wrong. For this practical attitude, any metaphysical attitude regarding the moral status will not matter. Legal theory puts all metaphysical assumptions regarding the nature of the mind and its properties into brackets. I have been drawing on the legal analysis because it is able to express very clearly how our categorizing attitude is going to be transformed by this sort of technological intervention which makes it clear that it is not the differences between machine and human that matter. The symbiosis indeed highlights a space-in-between that allows human beings to be involved in an effective seamless relationship with the machine.

According to the more recent developments in robotics, the robot is no longer a mere extension of the human, as in traditional teleoperation, rather the human being, will fully merge with the machine, creating a unit or symbiont whose physical and cognitive capability is more than the sum of those of the human and machine symbionts. These new technologies have the potential to collapse this ontological threshold into a single dimension while offering intimate experiences that are triggered by the interaction directly. By looking to legal theory, we may be able to better address issues regarding the conceptual analysis of the mind. This is the case since the symbiosis does affect how we conceptualize our experience and reshapes the status of the objects closely linked to us. As a result, this will also motivate an alteration in the way we will consider the artifact. While particular events may be explained by physical fact, this cannot justify a different interpretation of subjectivity and other humanlike features. Spelling out how technologies find their parasitic way to accomplish a new status is an intriguing opportunity to perform a kind

of thought-experiment. It leads to a sort of hybridization of the fully human character, blurring the distinctions between human and artifact. To be sure, major questions still stay open.

According to the analysis of the symbiosis, should we reverse the logical precedence of the experiential and normative aspects of the subjectivity and personality over the biological-machinic elements? Should we rely on a process of technical feasibility and usability in order to make sense of consciousness, personality, autonomy, and responsibility? Will self-objectification be the key feature that can make sense of the cultural dimension of human existence? Nonetheless, a consensus about the basic rudiments of the relationship between human and machine appears to be emerging. Hybridity is seen as a constitutive dynamics of techno-social civilization, which is linked with the dialectic of inclusion and exclusion of natural and artificial elements. Critics insist that we think of the notion of human in a different sense compared to its technological extensions, whereas their defenders describe new forms of human–machine interaction as indispensable fore-runners to more inclusive and advanced forms of self-understanding. What is certain is that all these questions are released from metaphysical burdens. The seamless relationship between human and machine, by modifying personality and responsibility attribution, may contribute to a larger inquiry as an additional epistemic device. This feature of the self-objectification dynamics can be clarified by contrasting it with Kapp's view, as the classic example of alienation theories of technology.

4. Alienation

There is a dialectic involved here, insofar as even if human enhancement technologies have their positive impact in a liberal way of thinking over human nature, these same technologies carry within themselves other possibilities, which are often of quite contrary quality. Therefore, it is misleading to take the positive possibilities, which we have recorded through the legal clarification, for granted. Technology is a constitutive element of the practices and institutions of the human condition. Technology constitutes an area of ethical investigation, and at the same time presents challenges to moral theory because it transforms reality and demands that moral theory adapt to a changed situation such as the one that accompanies the gradual process of human merging with the machine. I will critically discuss the conceptualization of technology in terms of mere instrument. My intention is to contend the partiality of this way of understanding the technology. In my view, this initial conceptualization needs to be complemented by a second perspective, which is focused on technology as a process of mediation that tends to transform what is at stake.

When one is about to decide on the transformation whose source is technology, controversies may arise about the ways in which the relationships between technology and society can be conceptualized and interpreted. The interpretation of the social and ethical implications also depends on the characterization that we will give of the technology. The conceptualization of the relationship between technology and human is responsible for various questions about nature, the human nature, human action, autonomy, freedom and much more. It is not the intention of this contribution to explore all these dimensions. I will focus only on the conceptions of technology even if some of the results of my argument will have some impact on the other dimensions as well.

Reflecting on the way technology has been conceptualized can help to clarify its transformative effects. The process of our merging with machines is certainly one of the most remarkable. A first form of conceptualization has been developed within the philosophical anthropology. According to this philosophical movement, technology was conceived as compensation for the biological deficits of the human being. Cassirer observed that at the basis of technical action there is a teleological relationship that presupposes a rational actor or a multitude of actors acting cooperatively [33]. This first part of the characterization is the most conservative and does not question the current ontology. To a certain extent, it matches the "introduction stage" described by Moor [1]. However, it is partial and fails to formulate an adequate representation of the phenomenon that it

wants to capture. There is, therefore, a second element, and that is culture. If we analyze technology from this point of view, then we will have to admit that it is part of creativity and freedom of spirit and therefore able to change the structure with which we are familiar. It is capable of having transformative effects such as transformation of the ontology. To a certain extent, Cassirer's cultural dimension matches the "permeation stage" described by Moor [1]. Drawing on this double characterization, further considerations have been derived. The first is that technology can be conceived as a tool that set us free from natural constraints and thus exonerates us, for example, from dull, dirty, and dangerous tasks.

However, technology can also be conceived as a risk of alienation of humans from themselves. In this interpretation, we place the profiles of dehumanization that have so much weight in current debates. In the philosophy of technology, there is the idea of conforming to the objective forms of human activity and therefore to attribute to them a function of mediation of self-understanding. By considering technology, it would be possible to know what becomes of us. Technology highlights the fact that the actions mediated by it always include a concrete intervention in the material environment and therefore always represents a form of a relationship with nature. Technology then is both a relationship with nature and, since the human being himself belongs to nature, it also includes a certain relationship with oneself and therefore also with the immaterial, symbolic and normative dimension. Human beings then not only find themselves in their artifacts, but above all they recognize themselves in them. This thesis allows Kapp to explain the human body in terms of an organism by using the objectification of the tools we create.

Kapp tries to explain the human body, its being organism through its self-made tools. As they are regarded as human unconscious alienation, human organs now appear to be extensions of artifacts, and these themselves can thus become models for their exploration and interpretation. Since artifacts and their functional context are meant to be subject to natural laws, it is clear that they provide a model for research on the human organism, provided that it too is subject to natural laws, and is therefore doomed to lose track of both the phenomenal character of consciousness and normative processes.

Kapp has focused on the interplay of reflections that exists between technology and the human. If the human is essentially technical, then technology is the ground of the human culture. In this way, then technology will not be limited to putting mechanical pieces side by side with human ones. In fact, the process of becoming technical will bring to the maximum development the use of biotechnological options by eroding the dominion of the natural that for a long time had remained beyond the possibilities of human influence. It will not be possible to hide that such a process of conquest of territories previously untouched by moral judgments has a potential impact on the theoretical and symbolic dimension. This is the typical shift that the philosophy of technology allows to implement with epistemic benefits. Thinking about human organs starting from artifacts, conceived as unconscious alienation of the human, organs now appear from this overturned perspective as extensions of artifacts. The objectification allowed by this epistemic deviation allows them to be considered models for their exploration and interpretation. There is, however, a caveat. Once the model for their understanding is that of artifacts, they will lose their inner face. Undoubtedly, artifacts and their functional context are constitutively subject to laws, it is clear that they can serve as a model for research on the human organism, provided, however, that it is also understood only in those configurations that are governed by natural laws. However, this epistemological move has an undoubted value. Following Ernst Kapp's, one might be convinced to appeal to his notion of alienation within his theory of technology as a self-knowledge device.

The central idea here is that, through technology, we may come across human features. Technology serves us as tool for knowing our constitution. It represents a reliable way to conceptualize us. My thesis is that while we can use such hybrid forms to explore the boundaries of our self-objectification, the direction of the human–machine interaction in the difficult ethical situations must however remain that of adapting the machine to the human. This mirroring dynamics exposes the transformations of the human vulnerability.

The alienation nexus includes vulnerability and direction. The normative indication is that the adaptation should be determined according to human and not to machinic standards. Overall, the risk of self-objectification, which becomes alienation, is very well identified by Kapp's theory. The recent attitude puts forward by engineers, computer scientist, and philosophers that contends that "buildings robot is philosophy making" [34,35] is somehow echoing Kapp's philosophy of technology. This means that the process of building robots forces us to reflect on what human capabilities are, and is therefore comparable to a process of self-knowledge. As the ancient Greeks relied on the Oracle of Delphi, so we can count on the construction of robots to obtain knowledge of ourselves. I will discuss this claim and in doing so I will borrow Ernst Kapp's terminology. In Kapp's account, the human being is to be found in her artifacts; in her technical culture, she recognizes herself there [14]. A traditional view says that there is no increment of self-understanding without mirroring of a mind-external world. So far there is no disagreement between Kapp [14] and Pfeiffer and Bongard [34] and Wallach and Allen [35]. With the exception that the Kapp's argument has two aspects, while the recent account restricts itself to the self-knowledge claim without further caveat.

On the one hand, Kapp argues for the epistemic potential of technology. On the other hand, he warns against the risks of sorting humans and things as items of similar shape and features. On Kapp's account these are the ethical boundaries of the self-objectivation.

This analysis may accommodate a number of worries connected to the image of the human being as a well-functioning machine. Therefore, we can address this tendency as the process of *mechanization of the human being*. For such reasons, the interventions in the brain are research topics and cannot serve as a normative tool. It would be wrong to use the insights from the process of self-objectivation to formulate ethical recommendations. Ethical recommendation concerning what does happen when humans share the same environment with highly automated systems can set the stage for a more integrated perspective on the human mind. As they identify the direction of the adjustment, they can also suggest that the project to understand the mind will root in self-reports and capacity to engage in normative practices and discourses. Experience and morality are to be moved out of the lab and to the street, office, kitchen, bar or wherever people happen to be when they feel, think, and act.

From a normative perspective, we can state that highly automated systems must be designed in a way that they must adapt more closely to the communication behavior of humans and not, conversely, demand increased adaptation performance from humans, who may not be willing to accomplish. If this adaptation goes the opposite path, we consider this—from an ethical perspective—as an undesirable development that may result in dehumanization processes. If the aim is to serve society, then innovation has to comply with human aims and desires. Otherwise, innovation without ethical dimension may result as a blind kind of innovation [36].

While the legal example shows how puzzling the hybrid constellation can be, the ethical analysis of human enhancement technologies is a source of hard moral questions. Intervening in the brain involves engineering issues and ethical problems, which can result in the conceptual and philosophical re-examination of our familiar insights.

5. Conclusions

The main objective of this paper was to discuss the ways in which human enhancement technologies such as Brain/Neural Computer Interfaces and robotics have a bearing on theoretical and methodological issues in philosophy. This is important, because the way we theorize and conceive of the mind's features during the process of building new technologies has an impact on normative discourse and practice. A deep understanding of how our mind perceives, thinks, and acts is not to be achieved just by means of one approach, which objectifies the mind. What we do need is an integrating perspective among the competing approaches. The feeling of being the owner of the action is just one successful area of investigation, which can benefit from different approaches and grow into a reliable

account of our sense of agency, which can be enhanced in the presence of technical artifacts interacting with the body and the environment. In this sense, it is important to protect this symbiosis and prevent, control and monitor misuse, dual-use, and other non-intended applications, so that the benefits are not outweighed by the damages. The exploitation of the technological artifacts, their moral status, and their ethical governance are able to provide new understanding on aspects of the mind, such as the notions of self, agency, and responsibility.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Moor, J.H. WHAT IS COMPUTER ETHICS? *Metaphilosophy* **1985**, *16*, 266–275. [CrossRef]
2. European Group on Ethics in Science and New Technologies. *Ethical Aspects of ICT Implants in the Human Body*; EGE: Brussels, Belgium, 2005; p. 20.
3. Merkel, R.; Boer, G.; Fegert, J.; Galert, T.; Hartmann, D.; Nuttin, B.; Rosahl, S. *Intervening in the Brain. Changing Psyche and Society*; Springer: Berlin/Heidelberg, Germany, 2007.
4. Lucivero, F.; Tamburrini, G. Ethical monitoring of brain-machine interfaces. *AI Soc.* **2007**, *22*, 449–460. [CrossRef]
5. Steinert, S.; Friedrich, O. Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces. *Sci. Eng. Ethic.* **2020**, *26*, 351–367. [CrossRef] [PubMed]
6. Battaglia, F. The embodied self and the feeling of being alive. In *The Feeling of Being Alive*; Marienberg, S., Fingerhut, J., Eds.; Walter de Gruyter: Berlin, Germany; New York, NY, USA, 2012; pp. 201–221.
7. Battaglia, F.; Carnevale, A. Epistemological and Moral Problems with Human Enhancement. *J. Philos. Stud.* **2014**, *7*, 3–20. Available online: http://www.humanamente.eu/index.php/HM/article/view/111 (accessed on 18 January 2021).
8. Straub, J.; Sieben, A.; Sabisch-Fechtelpeter, K. Menschen besser machen. In *Menschen Machen. Die Hellen Und Die Dunklen SEITEN Humanwissenschftlicher Optimierungsprogramme*; Sieben, A., Sabisch-Fechtelpeter, K., Straub, J., Eds.; Transcript: Bielefeld, Germany, 2012; pp. 27–75.
9. Coeckelbergh, M. Drones, information technology, and distance: Mapping the moral epistemology of remote fighting. *Ethic- Inf. Technol.* **2013**, *15*, 87–98. [CrossRef]
10. Cockelbergh, M. *Human Being @Risk*; Springer: Berlin/Heidelberg, Germany, 2013.
11. Parens, E. (Ed.) *Enhancing Human Traits: Ethical and Social Implications*; Georgetown University Press: Washington, DC, USA, 1998.
12. Savulescu, J.; Bostrom, N. *Human Enhancement*; Informa UK Limited: London, UK, 2019; pp. 319–334.
13. Savulescu, J.; ter Meulen, R.; Kahane, G. (Eds.) *Enhancing Human Capacities*; Wiley-Blackwell: New York, NY, USA, 2011.
14. Kapp, E. Selections from Elements of a Philosophy of Technology. *Grey Room* **2018**, *72*, 16–35. [CrossRef]
15. Nagel, T. *Mind and Cosmos*; Oxford University Press: Oxford, UK, 2012.
16. Chalmers, D.J. *The Character of Consciousness*; Oxford University Press (OUP): Oxford, UK, 2010.
17. Jackson, F. Consciousness. In *The Oxford Handbook of Contemporary Philosophy*; Jackson, F., Smith, M., Eds.; Oxford University Press: New York, NY, USA, 2009; pp. 363–390.
18. Gallagher, S.; Zahavi, D. *The Phenomenological Mind*; Routledge: London, UK, 2020.
19. Darwall, S. *The Second-Person Standpoint*; JSTOR: New York, NY, USA, 2009.
20. Korsgaard, C.M.; O'Neill, O. *The Sources of Normativity*; Cambridge University Press (CUP): Cambridge, UK, 1996.
21. Scanlon, T.M. *What We Owe to Each Other*; Indiana University Press: Bloomington, IN, USA, 2000.
22. Craver, C.F. *Explaining the Brain*; Oxford University Press (OUP): Oxford, UK, 2007.
23. Jackson, F. What Mary Didn't Know. *J. Philos.* **1986**, *83*, 291. [CrossRef]
24. Davidson, D. Mental Events. In *Essays on Actions and Events*; Clarendon Press: Oxford, UK, 1980; pp. 207–224.
25. Haggard, P. Conscious intention and motor cognition. *Trends Cogn. Sci.* **2005**, *9*, 290–295. [CrossRef] [PubMed]
26. Gallagher, S. The Natural Philosophy of Agency. *Philos. Compass* **2007**, *2*, 347–357. [CrossRef]
27. Synofzik, M.; Vosgerau, G.; Newen, A. Beyond the comparator model: A multifactorial two-step account of agency. *Conscious Cogn.* **2008**, *17*, 219–239. [CrossRef] [PubMed]
28. Santello, M.; Bianchi, M.; Gabiccini, M.; Ricciardi, E.; Salvietti, G.; Prattichizzo, D.; Ernst, M.; Moscatelli, A.; Jörntell, H.; Kappers, A.M.; et al. Hand synergies: Integration of robotics and neuroscience for understanding the control of biological and artificial hands. *Phys. Life Rev.* **2016**, *17*, 1–23. [CrossRef] [PubMed]
29. Miller, L.E.; Fabio, C.; Ravenda, V.; Bahmad, S.; Koun, E.; Salemme, R.; Luauté, J.; Bolognini, N.; Hayward, V.; Farnè, A. Somatosensory Cortex Efficiently Processes Touch Located Beyond the Body. *Curr. Biol.* **2019**, *29*, 4276–4283.e5. [CrossRef] [PubMed]
30. Crane, T. *The Mechanical Mind*; Routledge: London, UK, 2017.

31. Kersten, J. Die maschinelle Person—Neue Regeln für den Maschinenpark? In *Roboter, Computer und Hybride. Was Ereignet Sich Zwischen Menschen und Maschinen*; TTN-Studien—Schriften aus dem Institut Technik—Theologie—Naturwissenschaften; Manzeschke, A., Karsch, F., Eds.; TTN-Studien: Baden-Baden, Germany, 2016; Volume 5, pp. 89–101.

32. Sloterdijk, P. Kränkung durch Maschinen. Zur Epochenbedeutung der neuesten Medizintechnologie. In *Nicht gerettet. Versuche nach Heidegger*; Sloterdijk, P., Ed.; Suhrkamp: Frankfurt, Germany; pp. 338–366.

33. Cassirer, E. Form and Technology. In *Ernst Cassirer on Form and Technology, Contemporary Readings*; Hoel, A., Folkvord, I., Eds.; Palgrave Macmillan: London, UK, 2012; pp. 15–53.

34. Pfeifer, R.; Bongard, J. *How the Body Shapes the Way We Think: A New View of Intelligence*; MIT Press: Cambridge, MA, USA, 2007.

35. Wallach, W.; Allen, C. *Moral Machines: Teaching Robots Right from Wrong*; Oxford University Press: Oxford, UK, 2009.

36. Fabris, A. *Ethics of Information and Communication Technologies*; Springer International Publishing: Heidelberg, Germany, 2018.