Organon F 25 (4) 2018: 517-538

Verbeek on the Moral Agency of Artifacts

EHSAN ARZROOMCHILAR – DANIEL D. NOVOTNÝ[1]

ABSTRACT: One of the important questions discussed by philosophers of technology has to do with the moral significance of artifacts in human life. While many philosophers agree that artifacts do have moral significance attached to them, opinions vary as to how it is to be construed. In this paper we deal with the approach of the influential Dutch philosopher of technology Peter Paul Verbeek. He criticizes traditional ethical theories for assuming that whatever relevancy artifacts have for morality is entirely dependent on human beings, since artifacts are mere passive instruments of human agency. In contrast, he develops a view of moral agency that includes artifacts and that ascribes moral agency to human-technology hybrids rather than to humans as such. The goal of this paper is to elucidate Verbeek's account of moral agency and evaluate it. We also deal with his views on postphenomenology and mediation underlying this account. Although the general gist of our paper is expository, we point out to several problems for Verbeek's account.

KEYWORDS: Artifacts – mediation – moral agency – Peter Paul Verbeek – postphenomenology – technology.

[1] Received: 27 April 2018 / Accepted: 13 September 2018

✉ Ehsan Arzroomchilar
Faculty of Theology, University of South Bohemia in České Budějovice
Kněžská 8, 370 01 České Budějovice, Czech Republic
e-mail: ehsan.arzeroomchilar@gmail

✉ Daniel D. Novotný
Faculty of Theology, University of South Bohemia in České Budějovice
Kněžská 8, 370 01 České Budějovice, Czech Republic
e-mail: danieldnovotny@gmail.com

1. Introduction

Cars and trains, microwaves and refrigerators, TV sets and mobile phones, pencils, cups and glasses … countless artifacts from simple tools to sophisticated devices are ubiquitous in our lives. And there are many ontological, epistemological, ethical and other questions that philosophers may ask about these and other technological items. One of the main issues discussed by philosophers of technology today has to do with their moral significance. While many agree that technology and artifacts have moral significance attached to them, there are diverse views on how to construe it. In this paper we deal with the approach of the influential Dutch philosopher of technology Peter Paul Verbeek (University of Twente). In particular we deal with his account of moral agency, which is crucial for his project of reassessing the moral significance of artifacts.[2] Verbeek criticizes traditional ethical theories for wrongly assuming that whatever relevancy artifacts may have for morality is entirely dependent on human beings, since artifacts are mere passive instruments of human agency. In contrast, he develops a view of moral agency that includes artifacts and in which moral agency is ascribed to human-technology hybrids rather than to humans as such.

In what follows we first locate Verbeek's approach to the moral agency of artifacts in the broader context of contemporary philosophical studies related to technology (Section 2). Then we deal with postphenomenology, the philosophical background on which Verbeek draws (Section 3). Next we discuss one of his central concepts, namely that of mediation (Section 4), in order to better understand his view of moral agency (Section 5).

---

[2] We mostly draw on his *Moralizing Technology – Understanding and Designing the Morality of Things* (2011), where he develops his theory of moral agency most fully. We also take into account his other writings, especially his first book *What Things Do: Philosophical Reflections on Technology, Agency, and Design* (2005). We focus specifically on Verbeek's account of moral agency as we could not identify explicit discussion of agency in general. In order to facilitate a better understanding we occasionally provide direct references to authors that exerted great influence on him, namely Don Ihde and Bruno Latour. In doing this we by no means aspire to be exhaustive as our focus is Verbeek's theory as such, not an assessment of his reception of other authors. (We omit, for instance, references to Maurice Merleau-Ponty, Michel Foucault, Albert Borgman, Peter Sloterdijk and others.)

Finally, we evaluate his view and point out to some of its difficulties (Section 6).

2. Context

Verbeek's approach to moral agency of artifacts may be usefully related to two key questions that are commonly posed in philosophy of technology: (1) To what extent do humans shape technological products and processes? And (2) in what ways do technological products and processes shape human actions and experiences? (Mitcham & Waelbers, 2009).

As Mitcham and Waelbers have pointed out, in response to the first question we could imagine a spectrum wherein *voluntarism* is located at one end and *determinism* at the other. The advocates of voluntarism hold that the development of technologies is determined exclusively by human will, and that technological development is therefore malleable. Advocates of determinism, on the contrary, believe that technological development is determined by the internal logic of technologies themselves. Ellul, for instance, holds that old technologies are automatically replaced by those which are more efficient (Ellul, 1964; cf. Verbeek 2005, 11).

In response to the second question, we could propose another spectrum wherein *instrumentalism* is located at one extreme and *substantivism* at the other. Considering the relation between humans and technology, instrumentalists hold that technology is humanity's slave and thus it is nothing but an instrument in human hands. Advocates of substantivism, on the other hand, regard this relation as reverse and believe that technology is something "substantial", dominates over humanity and indeed holds us in its clutches.[3]

Now Verbeek's account of moral agency of artifacts is clearly neither voluntarist nor instrumentalist. Technologies are not determined exclusively by human will and they are not mere instruments. Relations between

[3] While these views are located at the two ends of a spectrum, adopting intermediary positions is also possible. For example, one can adopt a view in which neither humanity nor technology holds the other in its power. In such a perspective, the development of technology is due neither to human decisions alone, nor exclusively to the internal logic of technologies themselves. (Verbeek's view is a version of intermediary position).

humans and "things" are more complex. Already in *What Things Do: Philosophical Reflections on Technology, Agency, and Design*, first published in Dutch in 2000, Verbeek criticizes one-sided technophobia of earlier thinkers such as Heidegger and Jaspers that obscured intertwined, mediated character of human involvement with technology. He has been also increasingly interested in exploring the moral dimension of human-technology relations (Verbeek 2005, 212ff. and 2011). Drawing on earlier authors Verbeek approvingly acknowledges Hans Achterhuis's call for "moralization of devices" and Bruno Latour's view that "morality is not only to be found in humans but also in things" (Verbeek 2011, viii). As we shall see below, Verbeek provides various examples to show that "nonhuman entities are bursting with morality" (Verbeek 2011, 2).

Verbeek moreover vehemently opposes the "[m]ainstream ethical theory [that] does not leave much room for … a moral dimension of material objects" (Verbeek 2011, 2). What is the reason for this neglect? According to Verbeek it is the mistaken assumption that "technologies lack consciousness, rationality, freedom, and intentionality" and hence "morality ... is a solely human affair" (Verbeek 2011, 6). In order to amend this situation Verbeek sets out to develop "a notion of moral agency that does include material entities" (Verbeek 2011, 18).[4]

The above given quotes may seem to suggest that Verbeek sympathizes with determinism and substantivism in that artifacts are the main bearers of moral agency. However, although artifacts actively cooperate in shaping human experiences and actions we are not *completely* in their clutches. In fact, "ethics should be approached as a matter of human-technological association" (Verbeek 2011, 13). This means that "rather than separating or purifying 'humans and nonhumans' … the ethics of technology needs to hybridize them" (Verbeek 2011, 14). In other words, we cannot "hold on to the autonomy of the human subject as a prerequisite for moral agency; rather we need to replace the 'prime mover' status of the human subject with technologically mediated intentions" (Verbeek 2011, 16). In this way we get past the "subject-object distinction" and "articulate an 'amodern' perspective on ethics in which moral agency becomes a matter of human-

[4] Albeit Verbeek adds a cautious qualification here: moral agency of material entities "at the same time recognizes and articulates the differences between human and nonhuman elements of moral agency" (Verbeek 2011, 18).

technology hybrids rather than an exclusively human affair" (Verbeek 2011, 17).

We see that in his discussions of technology Verbeek refuses to separate humans from artifacts and hence his approach is best characterized as a peculiar intermediary position which is neither voluntarist nor determinist, neither instrumentalist nor substantivist. In order to understand Verbeek's approach to the moral agency of artifacts better let us now review what he says about "postphenomenology", the philosophical framework within which he develops his views.

3. Postphenomenology

What is postphenomenology? In the words of the American philosopher Don Ihde, its major proponent and initiator, it is "a modified, hybrid phenomenology" that avoids "the problems and misunderstandings of phenomenology as a subjectivist philosophy, sometimes taken as antiscientific" (Ihde 2009). Verbeek also understands it as "a new interpretation of the phenomenological tradition" but gives it "a broader definition" than Ihde (Verbeek 2005, 101). Verbeek's postphenomenology weaves together three streams: early philosophy of technology, phenomenology, and the empirical turn in technological studies from the 1980s (Verbeek 2005). Let us deal with them in turn.

3.1. Early philosophy of technology

Artifacts and technology have been the subject of occasional philosophical reflection since Plato and Aristotle (Schummer 2001). However, the urgency of the "problem of technology"[5] became obvious only in modern times. It was only in mid-nineteenth century Germany that sustained philosophical effort to understand technology led to the constitution of

[5] Peter Kwasniewski usefully characterizes the problem of technology in the context of his discussion of Leibniz as follows: "Leibniz's contribution to what may be called 'the problem of technology' (in the original sense of *techne* – art, craft, invention …) serves to highlight a major tension between belief in a fixed natural order providentially arranged for the best, and belief in a world of infinite possibilities, malleable to human hands and subject to human minds" (Kwasniewski 2017, 116).

a philosophical sub-field called "philosophy of technology" (Raydon) (Franssen et al. 2015). Most early contributors to the field lived and worked in Germany. Technology was a prominent theme in philosophers such as Karl Jaspers, Martin Heidegger, Hans Jonas, and others.[6]

For the most part Verbeek views the earlier, German phase of the philosophy of technology as a dialectics between Jaspers's existentialism and Heidegger's hermeneutics (Jonas's contribution is not explicitly taken into account). Whereas Jaspers was looking for answers to questions about the role that technology plays in human existence and action, Heidegger was searching for the way reality is disclosed by technology (Verbeek 2005, 16; Heidegger 1954/1977; Jaspers 1931/1951, 1958/1963). Verbeek takes the questions asked by Jaspers and Heidegger to be of crucial importance but finds their answers flawed. The most serious problem, Verbeek believes, is that they deal with the subject from a "transcendental perspective" (Verbeek 2005, 100). The transcendental perspective, as Verbeek understands it, is a perspective in which the researcher, instead of dealing with technological artifacts themselves and instead of examining their effects and consequences in daily life, addresses the "origins" of technology and the forces shaping it. For instance, in Jaspers's view technology is the fruit of a special functional way of looking at the world (Verbeek 2005, 28-30), while Heidegger holds that technology is the revealing of reality as a "standing-reserve of raw material" (Verbeek 2005, 95; Heidegger 1977, 10). Against this Verbeek calls for the study of technology which takes the particular artifacts themselves as its point of departure. This departure ought to issue in an evaluation of the outcomes engendered by those artifacts. We should not start with the artifacts and end with the causes and grounds of their formation. The direction of research should be exactly the reverse of the one found in the works of Jaspers and Heidegger.

In sum, what postphenomenology borrows from the early philosophy of technology are the questions: What kinds of impact does technology

---

[6] This earlier tradition of philosophy of technology, called now "humanities philosophy of technology" for its continuity with humanities and social sciences, has been complemented more recently by "analytical philosophy of technology", which is more closely related to philosophy of science and analytical philosophy (Franssen et al., 2015). For the most part Verbeek engages only the humanities philosophy of technology.

have on the existence and actions of humanity? How does technology affect the human experience of existence, and how does it disclose the world to us?

3.2. Phenomenology

The second major ingredient of postphenomenology is phenomenology. Phenomenology became established in the early twentieth century in the works of Edmund Husserl and authors influenced by him. It may be usefully characterized as "the study of structures of consciousness as experienced from the first-person point of view" (Smith 2013).

Phenomenology starts from phenomena as they appear in consciousness. In addition to phenomena, however, consciousness is also given and it is viewed as the place where humans and the world meet. The world is constructed/constituted in consciousness. And consciousness is always consciousness of something in the world. Therefore neither phenomena of the world nor consciousness can be found without the other. Human consciousness and the world can bear meaning only in relation with one another (Verbeek 2005, 109-112). It is this last point that is so attractive to Verbeek.

Verbeek and other postphenomenologists retain much sympathy for phenomenology. This is firstly because phenomenology's point of departure is to go back to the "things themselves". Indeed, postphenomenologists are interested in the study of technology from the artifacts themselves. Secondly, they agree with the phenomenological claim that the empirical sciences only represent *one* aspect of reality, not the fullness of reality as such. Nevertheless, despite acknowledging that our views of reality are always aspectual, phenomenologists assume that by employing their method they do get to represent "the true original world" in the end. Verbeek rejects this assumption (Verbeek 2005, 105). If phenomenology is to be used in studying technology, some modifications will be required. Verbeek makes the following two (Verbeek 2005, 104-116):

1. The possibility of any kind of access to the "original world" should be abandoned. Every sort of encounter with the world is "relative", not in the sense of an epistemological relativism, but rather in a more literal sense of the analysis of relations. It is the best approach one

can adopt, for "subject and object are not merely intertwined with each other but constitute each other" (Verbeek 2005, 112).

2. Phenomenology should be broadened to deal with any kind of relation between humanity and the world, including the relations that we are not conscious of. The relation between humanity and the world is not limited to the relations between conscious subjects and objects of which the subjects are conscious. The world is not just the source of *cognition* for humanity but, more significantly, it is where we *live*.

In sum, postphenomenology employs the phenomenological method, which is, however, modified in important ways. Most saliently it abandons aspirations to get to the "original world", emphases the interrelatedness of all experienced items and moves beyond what is presented in consciousness.

3.3. The empirical turn

The third main influence on Verbeek's postphenomenology has to do with a certain kind of empirical studies of technology emerging in the 1980s. Verbeek refers to Langdon Winner, the American scholar who discussed the low-hanging overpasses on Long Island in New York in "Do Artifacts Have Politics?" (Winner 1980). Those overpasses were deliberately built very low in order to prevent buses from using the road and allow only automobiles to pass underneath. At the time when these bridges were built this meant that racial minorities and the poor, who could not afford cars and who generally relied on public transportation, were effectively prevented from reaching the beaches. These overpasses shaped the ethnic and racial composition of people at the beach. What is remarkable about this investigation is its strongly empirical manner.

Another thinker acknowledged by Verbeek is the French anthropologist and philosopher Bruno Latour. He established a new framework by developing the so-called *actor-network* theory, studying artifacts as interactions between humans and artifacts. According to him humans and nonhumans are located in a network in a similar and indistinct way, with each component of that network cooperating. It is only the whole that acts. For example, in murdering someone with a gun, both the shooter and the gun are

responsible for that killing. This act is the result of the cooperation of the two "actants". Neither is able to carry out the task without the other (Verbeek 2005, 102). But while admiring the empirical character of such approaches to artifacts, Verbeek points out that we must not forget about the questions posed by the classical philosophy of technology (Verbeek 2005, 100).

Hence we see that Verbeek's version of postphenomenology pursues questions of earlier philosophy of technology by means of a modified phenomenology and a (certain version of) empirical studies pursued by Latour and others. Verbeek also stresses that postphenomenology should study both the hermeneutic and the existential aspects of artifact-human relations. It should deal with artifacts from the point of view of the role they might play in human perceptions of the world and how reality is thereby unfolded for humanity (the hermeneutic aspect). It should also deal with artifacts from the perspective of the role they might play in actions, behaviours, and in human existence generally (the existential aspect).

Having elucidated Verbeek's framework for studying technology we now turn to his views on mediation in order to interpret his account of the moral agency of artifacts.

4. Mediation

Artifacts are not just simple tools needed to attain human goals, their natures are not neutral. They have tremendous impact, sometimes foreseen and intended, at other times undesirable. Their influence in the world may even be unexpected and no one need be aware of it. Perhaps most importantly, however, artifacts at present *mediate* almost all our actions and perceptions. For instance, by sharing news and pictures of problems in developing countries, communication technologies have encouraged people in developed countries to spend more time and more money on charities (Waelbers 2011, 1). Another example: in some North European countries the length of the tube of the average vacuum cleaner is very short and thus is uncomfortable for men to work with, causing them back pain. This disinclines men to clean their homes and so reinforces sexist assumptions about who does housework (Waelbers 2011, 2).

Verbeek pursues the topic of mediation along the two lines of inquiry described in the previous section – hermeneutic and existential.

4.1. Hermeneutic aspects of mediation

Verbeek situates his account of the hermeneutic aspects of the mediation of artifacts into the context of the distinction between microperception and macroperception (Verbeek 2005, 122-123). The former concerns ordinary perception, such as seeing a tree or smelling a flower, while the latter concerns the framework within which sensory perceptions become meaningful. Artifacts have a huge impact on and hence a mediating role for both.

In their mediating role on the micro-level, artifacts lead to an amplification of some aspects and a reduction of others. For example, a thermographic camera shows some aspects of reality that we could never see without such equipment. At the same time some aspects of reality (including some non-visual dimensions) are reduced and distorted (Ihde 1991, 73-74).

Artifacts mediate human perceptions on the macro-level too. By changing the frameworks in which human interpretation occurs, they change our experiences. Postphenomenologists regard two of these frameworks as most important: cultural and scientific.

The cultural framework of macro-level mediation may be seen, for instance, in the appearance of communicative technologies and connections between different cultures. We are now prompted to see everything through different lenses, and so (typically) to show more tolerance toward different perspectives (Ihde 1993a). Also, thanks to the emerging modern technologies, humans are now compelled to make more choices; thus technologies create a "decisional burden" for modern humanity. This is obvious, for instance, in the advent of biomedical technology that forces people to make choices in situations that had traditionally been determined, such as when a foetus suffers from a serious disease detectable by our screening methods. We now have to choose between killing or letting "it" live. In Ihde's words, "The one choice I do not have is the choice not to make a choice" (Ihde 1990, 181).

The scientific framework of macro-level mediation concerns tools and equipment used in the constitution of scientific knowledge. This framework becomes ever more prominent as science plays an increasingly strong role in shaping the ways in which we interpret our world. We even evaluate our quotidian physical and mental well-being in medical and scientific terms. And it is also clear that scientific achievements are closely related to the advancement of our tools and equipment. Radio telescopes, for

instance, make things that are not accessible to the naked eye "perceivable". These mediated perceptions reveal entities that we would never have come to know without our mediating technologies. Technological instruments play an essential role in the generation of scientific knowledge, and studying this role is crucial for a proper understanding of contemporary science.

Artifacts, therefore, through altering human perception on micro- and macro-level play an undeniable role in how reality is revealed. Accordingly, postphenomenologists expand the notion of hermeneutics to apply it not only to texts but also to instruments and technological artifacts and their mediating roles. Hermeneutics in their hands becomes a kind of interpretation of objects (Ihde 1998, 139).

4.2. Existential aspects of mediation

Verbeek examines the mediation of artifacts from yet another angle, drawing on Latour's views. The core of his actor-network theory can be rendered as follows: Consider person A who murdered person B using a gun and aiming to take revenge. Latour's claim is that in such a situation we could not attribute this murdering only to the person who shoots the bullet, namely A; rather, the gun itself plays a mediating role in this event. In a Latourian perspective the scenario is this: A's "program of action" is taking revenge on B. On the other hand, the gun's program of action is shooting (not necessarily shooting a specific person). Out of these two programs of action a new third one arises, which is killing someone. This latter program of action is neither merely a result of A's program of action, nor exclusively a result of the gun's program of action; it is the outcome of a "composition" of both (Latour 1999). Latour sees all the actors, whether human or nonhuman, within a network in which they are constantly altering each other's program of action, resulting in a new program of action. In his words, "you are different with a gun in your hand; the gun is different with you holding it. You are another subject because you hold the gun; the gun is another object because it has entered into a relationship with you" (Latour 2005, 179-180). In his eyes, we should never see an artifact as a simple tool; rather we should consider it as having an agency analogically comparable to the agency of humans. "It will become more and more difficult to trace the border between the empire of the human and the realm of

technologies" (Latour 1992, 248). Artifacts thus continually shape our actions and deeds.

Latour's views are an important source for Verbeek who uses them to show that artifacts mediate both "how the world is present for human beings" (hermeneutic aspects of mediation) and how human beings are present in the world (existential aspects of mediation) (Verbeek 2005, 172). Verbeek also takes seriously Latour's thesis concerning the inseparability of humans and nonhumans.

5. Moral agency

Now, what do these points about postphenomenology and mediation have to do with the ethics and morality of artifacts? Since artifacts affect perceptions and human actions, they also affect morality. For instance, as we have seen, morality in the pathological form of racism became inbuilt into New York's overpasses. And Verbeek offers many other examples. Sonography is a method of examining foetuses by ultrasound. Consider a situation in which a pregnant woman finds, through sonography, that there is a high degree of probability that the child will be born with Down's syndrome. The finding may prompt the woman to abort the child and at any rate it will force her to decide what to do. We may also notice that the sonogram produces new meanings of what "foetus", "father", "mother" are and how they relate. It represents the foetus in specific ways: as being the same size as a newborn baby, in spite of the fact that it is much smaller, and as an independent entity freely floating in space, although it is closely linked to the mother. The sonogram isolates for us the experience of the foetus and separates it from its context, the mother (Boucher 2004, 12; Verbeek 2011, 24-25).

The sonogram also mediates the role of the mother and the father: the father, who formerly played an unimportant role, has now become a determining character in the new situation, deciding whether or not to abort. The mother, too, who formerly merely carried the baby, now, thanks to this instrument, has not only become a determining character but also a person whose uterus counts as posing a threat to the baby. On the other hand, depicting the foetus probably cements the bond between the mother, the father and the baby sooner than would happen without the technology and

may give the parents a feeling of being closer and more attached to the unborn child. In this way the mother and the father experience a new situation before the day of birth. But more than anything else, the aspect which makes the moral dimensions of this issue clearer is that with the sonogram the parents are converted from "expecting parents" to "deciding parents" (Verbeek 2011, 25-27). The sonogram brings about an unprecedented situation in which we could prevent the baby being born in the case of some kind of possible danger. The emergence of such dilemmas proves the ability of artifacts to create completely new moral situations. By mediating human perceptions and actions artifacts construct situations and objects, cooperating actively in depicting reality.

While the moral significance of artifacts is generally acknowledged, they are not commonly regarded as a substantial part of the moral sphere. Rather, they are seen as neutral and more or less unimportant tools (Smith 2003, 183). By contrast, according to Verbeek following Latour, the moral agency of artifacts is similar and comparable to that of humans (Waelbers 2001, 31-33). Verbeek, however, is aware of the difficulties facing those who make such claims. Moral agents should possess intentionality and freedom, they should be able to form intentions and realize them (Verbeek 2011, 54). Ordinarily, however, we do not think of artifacts as being intentional and free and hence Verbeek needs to explain to us in what way he thinks they are.

Let us begin with Verbeek's views on intentionality. We speak of intentionality in two senses – as the ability to form intentions, and as a kind of directedness (Verbeek 2011, 55). In Verbeek's view these two senses are related in a similar way as the hermeneutic and existential dimensions of the mediation of artifacts. The former concerns our perceptions of reality, whereas the latter our activities in reality (Verbeek 2011, 55-56). This means that our relations with the world are ordinarily mediated by artifacts. We either make contact with the world through artifacts, such as seeing the world through glasses, or technological artifacts shape our relation to reality as we make contact with the world, albeit remaining in the background, like the thermostats that automatically switch the heat on and off without our intervention. In all these kinds of relations, our intentionality is mediated by an artifact. Whenever we see beautiful scenery using binoculars, the intentionality is not just due to the human element, but seemingly due to a "human-artifact dyad". In other words, since the connection between

us and the world is not shaped just by our humanity alone, but rather with the help of artifacts, intentionality is not just a property of a human being but of a human-being-plus-artifact (Verbeek 2011, 56-58). As Verbeek puts it, "intentionality is distributed among the human and the nonhuman elements … [r]ather than being derived from human agents, [it] comes about in associations between humans and nonhumans" (Verbeek 2011, 58).

Now, Verbeek treats freedom in a similar way. Since artifacts don't have minds, it would seem inappropriate to ascribe freedom to them. However, artifacts often lead to completely unexpected outcomes as if they had their own minds. For example, energy-saving light bulbs were first used to decrease energy consumption, and are ordinarily cheap to run. As a consequence, people started to use those bulbs to light places that used to be dark (such as gardens), and eventually energy consumption increased. These and other examples lead Verbeek to think that artifacts should not be regarded as unfree. He also offers two more reasons to back up his claim. First, if we take freedom as an absolute concept, we could not count even humans as possessing it, since all people in all of their decisions are constrained by their era and the material environment and the artifacts to which they are related. So, to be free, it is sufficient to possess *some* degree of freedom (Verbeek 2011, 89). Such partial freedom is then also enjoyed by artifacts. Second, since human actions are mediated by artifacts, we should not think of human beings independently of their involvement with artifacts. Freedom is then a property of a human-artifact dyad. Hence, Verbeek concludes, freedom like intentionality is distributed among humans and artifacts (Verbeek 2011, 60).

We see that Verbeek has a way to ascribe intentionality and freedom to artifacts. Although, as we have seen, Verbeek appeals to various intuitions and considerations, his main reason for doing so has to do with his views of mediation and with his postphenomenological commitments. Human beings and artifacts co-constitute one another and form inseparable hybrids. It is these hybrids that are properly speaking intentional and free and hence moral agents. With respect to moral agency human beings are indistinguishable from artifacts. Both enter into the wholes that are moral agents in the proper sense, whereas taken alone they are not.

6. Problems

Within Verbeek's general postphenomenological commitments, as we have seen, it is not possible to ascribe moral agency to humans and at the same time to deny it to artifacts, despite Verbeek's occasional claims to the contrary. We take the general drift of his approach to blur any distinction between the moral agency of humans and artifacts.[7] In fact, Verbeek's postphenomenological understanding of (moral) agency does not seem to provide resources for drawing a clear distinction between humans and nonhumans – they are both parts of mutually constituted hybrids joined by mediation and other relations. He often insists on avoiding any kind of absolutizing subject and object (Verbeek 2005, 112). He also denies any gap between objectivity and subjectivity and speaks of mutual constitution. Artifacts (objects) and humans (subjects) are interwoven in such a way that they cannot be separated. In many passages Verbeek urges us to change our perspective on subjectivity and objectivity and, rather than assume them as pre-given, to consider them as co-shaped by one another (Verbeek 2005, 112). We, human beings, in some limited way do design and use artifacts, but they also structure our actions, perceptions and moralities. We stand in reciprocal relationships. We may initially decide to buy a car and use it, for instance, but immediately the car starts to affect our behaviour, expectations and thought. Once we have the car we may be able to rent a house far from our workplace, whereas without it we would have been obliged to live in the proximity of our workplace. Our behaviour has been affected by the fact that we own the car and as a result our situation within the world changes. There is no pure object vis-a-vis pure subject but all is

[7] At least at one occasion Verbeek claims that the idea that "technologies in themselves 'have' a form of agency that we normally only attribute to human beings" is a misreading of his work (Verbeek 2014, 79). He even notes that "it is in fact hard to find scholars who seriously defend the thesis that technologies can be full-blown moral agents just like human beings are" (Verbeek 2014, 79). We find these claims at odds with the general gist of his view. We hope to have made clear by now that Verbeek does not have resources to distinguish between the (moral) agency of artifacts and of humans. Also, by the way, it is not so rare to find scholars ascribing "full-blown agency" to some highly sophisticated artifacts such as AI robots, autonomous cars, etc. (These, however, are special subsets of artifacts, whereas Verbeek deals with artifacts in general).

"packed together" (Verbeek 2005, 164). The experiencing subject and the experienced object constitute one another.

Verbeek's claim that the subject is inseparable from the object allows him then to claim that "morality appears to be a coproduction of humans and nonhumans" (Verbeek 2014, 78) or that "morality is a hybrid affair" (Verbeek 2014, 80). One must overcome the view that morality is "located exclusively in humans" (Verbeek 2014, 80). The reason is simple – there is no pure human being, nor pure artifact.

The consequence of the human-artifact inseparability thesis is that human beings taken as such cannot be moral agents. Verbeek is aware that this calls for a new conception of moral agency. As he puts it: "rather than applying a human conception of agency to nonhumans, I rework the concept of agency in order to show that it should actually be seen as a property of hybrids rather than of humans only" (Verbeek 2009, 255). None of them could alone be deemed to be a self-standing agent. Morality is an attribute of a composite, of a network of human beings and artifacts.

There are three objections to Verbeek's view of the moral agency of artifacts that we would like to discuss.[8]

First, in our view Verbeek has misdescribed the moral status of artifacts by equalizing their contribution to moral acts. The conditions for an event to obtain ought not to be taken as a proper part of the event itself. Factors that bring about a specific framework within which a particular event happens are to be distinguished from the event itself. If I look at some beautiful scenery through a pair of binoculars, although this instrument does partly shape the framework of my experience, it is only *me* who is looking at that scenery, not me-*plus*-binoculars. The binoculars simply do not look at anything, they merely provide a framework within which I can see some things and not others. So even if it were appropriate to ascribe moral agency

[8] Other kinds of criticism have been put forward. Illis & Meijers (2014), for instance, object that Verbeek discusses only two necessary conditions of moral agency, intentionality and freedom, and ignores others. Philip Brey (2014) worries that by redefining moral agency and ascribing it to artifacts we are forced to ignore certain relevant features of human moral agents. Thorough and detailed criticism within the analytical tradition can be found in Peterson (2011) and (2017, 185); cf. also Selinger et al. (2012). While we are sympathetic to these kinds of criticism, our approach is more (although not exclusively) "internal", i.e. we point out to tensions within Verbeek's own philosophical commitments.

to artifact-human hybrids, it is humans rather than artifacts that are the primary locus of intentionality and freedom and hence of moral agency. The mediation of artifacts merely extends the sphere of moral agency which is grounded in and properly ascribed to human beings alone.

Why does Verbeek tend to obliterate distinctions between humans and artifacts? One of the reasons has to do with the way he describes his examples. True, no one had foreseen that the introduction of energy-saving bulbs would lead to an increase in energy consumption. This does not mean, however, that it was these light bulbs as such that decided that and hence are in the relevant sense responsible for it. We could have foreseen the danger and taken precautions. The light-bulbs could not. They are just what we make them to be. So, while it is true that artifacts dramatically change our lives and moralities and hence hardly are mere passive tools, they nevertheless do remain tools. It is to Verbeek's credit that he underscores the power of technology in our era and warns us about using and developing artifacts in an irresponsible way. However, we disagree with his account of the nature of artifacts and their moral agency. To highlight the role that technology can play in life one does not need to misrepresent the real functioning of artifacts.

Second, Verbeek's views on moral agency undo the distinction between artifacts and natural objects. If the only criterion that is at work in ascribing moral agency to things is whether it somehow affects the morality of actions, then (at least) some natural objects also qualify as moral co-agents. Hence we cannot distinguish them from artifacts. For it is clearly not just artifacts that structure our behaviours and steer our actions. Imagine, for instance, that Peter is walking in a dense forest and due to the existence of lots of trees and boughs he is obliged to constantly change direction. The trees and boughs act in the same way as a pair of binoculars does, except that they are natural objects, not artifacts. Does it make them moral agents as well? Is there any difference between the way that cars, knives or other artifacts affect our behaviour and that of the forest's effect? All of these put some specific restrictions on our activities, co-shaping our actions in a similar way. Or let's take another case. Suppose Mary runs into someone she hates and wants to take revenge on. Now imagine the following two possible scenarios. First, she takes a gun from her car and shoots the guy. He dies. Second, she leans over, picks up a big sharp stone and throws it at him. Again he dies. What is the difference? Both the stone and the gun

encouraged her to kill the guy and both shaped her actions. Stone-plus-Mary and gun-plus-Mary are both moral agents. Thus, Verbeek should acknowledge that (at least some) boughs and stones are moral co-agents. And if artifacts can be moral co-agents, then anything can. Perhaps Verbeek would be comfortable with this consequence. Many of us, however, would like to preserve the distinction between artifacts and natural objects and ascribe the status of moral co-agents only to some things.

Third, Verbeek has not left any place for the possibility of making a distinction between simple artifacts, such as a knife, and more evolved ones, such as autonomous cars. These are obviously not on the same level. For example, some of the more sophisticated artifacts may display abilities which make them more likely to qualify as moral agents than other simple ones. For a clearer grasp we can map out a spectrum representing various entities with respect to their intelligent behaviour dimension. In such a picture, we can locate natural objects at one extreme and human beings at the opposite one, with artifacts in between. It seems that not all artifacts could be situated at the same distance from humans. More complex artifacts, such as autonomous cars that need to "decide" how to react in unprecedented traffic situations, should be placed nearer to human beings than, for instance, knives. They imitate some aspects of human intelligent behaviour. Today's intelligent artifacts still lack some human abilities, such as moral deliberation or consciousness, but they do possess abilities such as learning, and (a sort of) thinking and decision-making. Perhaps eventually an AI robot will be constructed that will count as a full-blown moral agent. Simple artifacts such as flints or pencils, however, do not qualify. Our view, then, is that an adequate account of the morality of artifacts needs to do justice to the differences within their kind.[9]

Verbeek's remarks about the roles that artifacts can play in our lives are strikingly insightful. These observations should be taken seriously in designing and developing artifacts and in policy-making that concerns them. He has shed some light on how profoundly artifacts can change morality

[9] An anonymous referee points out that we assume in this objection that there is a hierarchy of the moral agency of artefacts, which may not be consistent with Verbeek's view about the "inseparability of humans and non-humans". We do not think so. And at any rate, even if our assumption is inconsistent is inconsistent with Verbeek's view, this only means that our objection is not internal but external.

and hence how important it is in applied and even in general ethics to take them into account. However, despite all of his contributions, the only lesson to take is that artifacts are much more powerful *tools* than we used to think, nothing less and nothing more. They are not as such agents nor co-agents, even though when we possess them there are lots of consequences for us humans. The ability of artifacts to change our lives requires us to become more careful and more responsible in developing and introducing technologies.

7. Conclusion

In this article we have dealt with Verbeek's view of the moral agency of artifacts. We have provided a broad philosophical background for his thinking by explaining the major elements of postphenomenology and the notion of mediation. We share Verbeek's sense of the urgency of "the problem of technology". A new technological invention usually profoundly modifies the moral situations that we have been facing so far. It is like "placing a drop of red dye into a beaker of clear water", to use Neil Postman's metaphor. After that we do not have clear water plus a spot of red dye but rather something entirely new (Postman 1998). Today ignoring the moral impact of artifacts is no longer an option. We value the contributions of Verbeek and other postphenomenologists to the ongoing debate about these issues. However, we have also found some aspects of his view, especially with respect to the moral agency of artifacts, wanting.

First, we have argued against placing artifacts and humans on the same level with respect to moral agency. In our view the Verbeekian approach by ascribing moral agency only to human-artefact hybrids runs the risk of anthropomorphizing artifacts and/or objectifying humans. Second, we have pointed out some undesirable consequences of Verbeek's views, namely the disappearance of the distinction between artifacts and natural objects. We think that philosophers of technology sensitive to phenomenological descriptions of our experiences should not abandon it. Third, we worry that Verbeek's claim that all human-artifact hybrids are moral agents hinders a proper understanding of the various levels at which some complex artifacts, such as robots or autonomous cars, may approximate moral agents while simple ones, such as binoculars or pens, do not.

We believe that Verbeek correctly shows that artifacts are not morally neutral, but in ascribing moral agency, albeit partial, to them he goes too far. His considerations undoubtedly show the need for responsibility in developing new technologies. However, technologies should not be assimilated to human moral agents. While artifacts profoundly affect morality, we cannot give up our own and proper individual responsibility as moral agents.

Acknowledgments

We would like to thank Martin Cajthaml, Světla Hanke Jarošová, Kateřina Kutarňová, Vojtěch Šimek, Peter Volek, participants of Bamberg-Budweis Philosophy Conference (Bamberg, July 4th, 2018), and especially to anonymous referees for discussion and comments. Novotný's work on this paper has been supported by Technology Agency of the Czech Republic (n. TL01000467 "Ethics of Autonomous Vehicles").

References

BOUCHER, J. (2004): Ultrasound – a Window to the Womb? Obstetric Ultrasound and the Abortion Rights Debate. *Journal of Medical Humanities* 25(1), 7-19.

BREY, P. (2014): From Moral Agents to Moral Factors: The Structural Ethics Approach. In: Kroes, P. & Verbeek, P. P. (eds.): *The Moral Status of Technical Artefacts*. Dordrecht: Springer.

ELLUL, J. (1964): *The Technological Society*. New York: Vintage Books.

FRANSSEN, M., LOKHORST, G.-J. & VAN DE POEL, I. (2015): Philosophy of Technology. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), URL = <https://plato.stanford.edu/archives/fall2015/entries/technology/>.

HEIDEGGER, M. (1954/1977): *The Question Concerning Technology and other Essays*. Trans. by William Lovitt. New York: Harper and Row.

IHDE, D. (1990): *Technology and the Lifeworld*. The Indiana Series in the Philosophy of Technology. Bloomington: Indiana University Press.

IHDE, D. (1991): *Instrumental Realism*. The Indiana Series in the Philosophy of Technology. Bloomington: Indiana University Press.

IHDE, D. (1993a): *Philosophy of Technology: An Introduction*. New York: Paragon House.

IHDE, D. (1998): *Expanding Hermeneutics*. Evanston: Northwestern University Press.

IHDE, D. (2009): *Postphenomenology and Technoscience.* Albany, NY: SUNY Press.

ILLIES, M. (2014): Artifacts, Agency, and Action Schemes. In: Kroes, P. & Verbeek, P. P. (eds.): *The Moral Status of Technical Artefacts*. Dordrecht: Springer.

JASPERS, K. (1931/1951): *Man in the Modern Age.* Trans. by Paul Eden and Paul Cedar. London: Routledge & Kegan Paul.

JASPERS, K. (1958/1963). *The Atom Bomb and the Future of Man.* Trans by E. B. Aston. Chicago: University of Chicago Press.

KROES, P. & VERBEEK, P. P. (eds.) (2014): *The Moral Status of Technical Artefacts*. Dordrecht: Springer.

KWASNIEWSKI, P. A. (2017): Divine Wisdom, Natural Order, and Human Intervention. Leibniz on the Intersection of Theology, Teleology, and Technology. *Studia Neoaristotelica* 14 (2), 115-138.

LATOUR, B. (1987). *Science in Action: How to Follow Scientists and Engineers through Society*. Milton Keynes: Open University Press.

LATOUR, B. (1992): Where Are the Missing Masses? The Sociology of a Few Mundane Artefacts. In: *Shaping Technology, Building Society*. Cambridge, MA: MIT Press.

LATOUR, B. (1999): *Pandora's Hope*. Cambridge, MA: Harvard University Press.

LATOUR, B. (2005): *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: Oxford University Press.

MITCHAM, C. & WAELBERS, K. (2009): Technology and Ethics: Overview. In: Berg Olsen, J., Pedersen, S. & Hendricks, V. (eds.): *A Companion to the Philosophy of Technology*. West Sussex: Wiley Blackwell, 367-383.

PETERSON, M. (2017): *The Ethics of Technology: A Geometric Analysis of Five Moral Principles*. Oxford: Oxford University Press.

PETERSON, M. & SPAHN, A. (2011): Can Technological Artefacts Be Moral Agents? *Science and Engineering Ethics* 17(3), 411-424.

POSTMAN, N. (1998): Five Things We Need to Know about Technological Change. Talk delivered in Denver Colorado March 28, 1998. URL = <http://web.cs.ucdavis.edu/~rogaway/classes/188/materials/postman.pdf>.

RAYDON, T. A. C.: Philosophy of Technology. *Internet Encyclopedia of Philosophy*, URL = <http://www.iep.utm.edu/technolo/>.

SCHMITT, R. (1967): Phenomenology. In: Edwards, P. (ed.): *Encyclopedia of Philosophy*. New York: Macmillan.

SCHUMMER, J. (2001): Aristotle on Technology and Nature. *Philosophia Naturalis* 38(1), 105-120.

SELINGER, E., IHDE, D., VAN DE POEL, I., PETERSON, M. & VERBEEK, P. P. (2012): Erratum to: Book Symposium on Peter Paul Verbeek's *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago: University of Chicago Press, 2011. *Philosophy & Technology* 25(4), 605-631.

SMITH, A. (2003): Do You Believe in Ethics? Latour and Ihde in the Trenches of the Sciences Wars. In: Ihde, D. & Selinger, E. (eds.): *Chasing Technoscience: Matrix for Materiality*. Bloomington: Indiana University Press, 182-194.

SMITH, D. W. (2013): Phenomenology. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), URL = <https://plato.stanford.edu/archives/sum2018/entries/phenomenology/>.

VERBEEK, P. P. (2005): *What Things Do: Philosophical Reflections on Technology, Agency and Design*. University Park, PA: Pennsylvania State University Press.

VERBEEK, P. P. (2009): Let's Make Things Better: A Reply to My Readers. *Human Studies* 32(2), 251-261.

VERBEEK, P. P. (2011): *Moralizing Technology: Understanding and Designing the Morality of Things*. London: University of Chicago Press.

WAELBERS, K. (2011): *Doing Good with Technologies: Taking Responsibility for the Social Role of Emerging Technologies*. New York: Springer.

WINNER, L. (1980): Do Artifacts Have Politics? *Daedalus* 109(1), 121-136.