



Technology and Neutrality

Sybre Heyndels¹

Received: 20 June 2023 / Accepted: 9 October 2023
© The Author(s) 2023

Abstract

This paper clarifies and answers the following question: is technology morally neutral? It is argued that the debate between proponents and opponents of the Neutrality Thesis depends on different underlying assumptions about the nature of technological artifacts. My central argument centres around the claim that a mere physicalistic vocabulary does not suffice in characterizing technological artifacts *as* artifacts, and that the concepts of *function* and *intention* are necessary to describe technological artifacts at the right level of description. Once this has been established, I demystify talk about the possible value-ladenness of technological artifacts by showing how these values can be empirically identified. I draw from examples in biology and the social sciences to show that there is a non-mysterious sense in which functions and values can be empirically identified. I conclude from this that technology can be value-laden and that its value-ladenness can both derive from the intended functions as well as the harmful non-intended functions of technological artifacts.

Keywords Technology · Neutrality Thesis · Moral values · Artifacts · Functions

1 Introduction

One of the philosophy of technology's central questions is the following: is technology morally neutral? Many philosophers argue that this is not the case. They usually do so by giving examples of technologies that seem to clearly embody moral values: death furnaces and gas chambers during the Holocaust (Katz, 2005), sea dikes and speed bumps (Kroes & van de Poel, 2014, 115), bridge designs which were intentionally created to make it harder to reach certain parts of the city by public transport (Winner, 1980), and city benches with anti-homeless spikes.¹ These philosophers

¹ The historical accuracy of Winner's example about Robert Moses' Parkway bridges has been disputed (Joerges, 1999; Woolgar & Cooper, 1999). Although it is rhetorically more convincing to use actual examples that are historically accurate, it is strictly speaking not necessary for counterexamples against the Neutrality Thesis to have been actualized. As I argue in Sect. 2, the Neutrality Thesis should be

✉ Sybre Heyndels
sybre.heyndels@kuleuven.be

¹ KU Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium

argue that there are technologies which are value-laden, because these technologies were designed with certain moral values or immoral purposes in mind. Hence, they conclude, technology is not morally neutral.

It is trickier to find explicit defences of the thesis that technology is morally neutral (hereafter the 'Neutrality Thesis'). One clear exception is Joseph Pitt, who elaborately argued for the Neutrality Thesis in his 2014 article "Guns Don't Kill, People Kill"; Values in and/or Around Technologies'. From the fact that the Neutrality Thesis is not often explicitly argued for by philosophers, however, it does not follow that the Neutrality Thesis lacks broader support. The Neutrality Thesis is often taken for granted and asserted as something obvious. Noam Chomsky reportedly claimed that technology "is basically neutral. It is like a hammer. The hammer doesn't care whether you use it to build a house or whether on torture, using it to crush someone's skull, the hammer can do either" (as quoted in Veletsianos, 2014). The British mathematician Hannah Fry wrote that "[n]o object or algorithm is ever either good or evil in itself. It's how they're used that matters. GPS was invented to launch nuclear missiles and now helps deliver pizzas. Pop music, played on repeat, has been deployed as a torture device" (Fry, 2018, 15). The Neutrality Thesis is sometimes also endorsed in discussions with professionals with a more technical background or among students pursuing an engineering, business, or computer science degree (Katz, 2005, 409). The claim that technology is morally neutral can be used by developers of new technologies to evade responsibility for the potentially bad consequences of these technologies. Therefore, the question concerning the Neutrality Thesis is not only of theoretical, but also of practical interest.

In this paper, I defend the thesis that technology can be value-laden, but that recent attempts at undermining the Neutrality Thesis remain unconvincing. In Section 2, I clarify what the Neutrality Thesis actually is. In Section 3, I look at one prominent argument advanced by Pitt (2014) in support of the Neutrality Thesis. The first premise states that a necessary condition for technological artifacts to be value-laden is that these moral values should be empirically identifiable. The second premise states that moral values are empirically unidentifiable. I evaluate a recent attempt to undermine the first premise of Pitt's argument by Miller (2021) and argue that Miller's criticism fails to show that this premise is false. In Section 4, I discuss the second premise of Pitt's argument. I reject Miller's recent attempt to undermine the truth of this premise, but I argue that this premise should be denied for other reasons. More specifically, I argue that Pitt's argument for (as well as certain implicit intuitions that seem to support) the Neutrality Thesis presupposes a concept of technological artifacts as objects that can be described in a merely physicalistic vocabulary. Against this, I argue that technological artifacts are not just 'mere' physical objects, but rather that they are *intentionally* created products with a particular *function* to do something. In other words, technological artifacts can only be properly described using a vocabulary that contains the concepts of *function*

Footnote 1 (continued)

understood as a claim about the necessary connection between technology and moral neutrality. Therefore, merely *possible* counterexamples in which technology is value-laden would, strictly speaking, suffice.

and *intention*. These are concepts that go beyond a (traditionally understood) physicalistic vocabulary. In Section 5, I develop my main argument in three steps. First, I distinguish between different types of functions that technological artifacts can have. Secondly, I argue about the sense in which it is possible to empirically identify these functions. Thirdly, I argue that the empirical identifiability of artifact functions leads to a natural way to defend the empirical identifiability of moral values from technological artifacts. This offers a straightforward way to reject Pitt's argument for the Neutrality Thesis. In Section 6, I provide further support for technology's possible value-ladenness by discussing the broader, 'sociotechnical' notion of technology. I end my discussion by drawing a connection between discussions of the neutrality of technology and the notion of responsibility.

2 The Neutrality Thesis

The first step towards evaluating the correctness of the Neutrality Thesis should be to clarify it. As a first approximation, the Neutrality Thesis can be captured as follows:

(NT) Technology is morally neutral.

This does not yet amount to a very informative characterization. Three elucidatory tasks need to be carried out here. First of all, the relevant notion of 'technology' should be clarified. Most recent discussions of the Neutrality Thesis (e.g. Pitt, 2014; Kroes & van de Poel, 2014; Miller, 2021; Klenk, 2021) focus on so-called 'technological artifacts'. However, there are also other uses of the word 'technology'. One use of 'technology' is to refer to a manipulation process, rather than to the end products of such a process. For example, 'gene technology' denotes the process of manipulating or removing genes. Another use of 'technology' is to refer to a "sociotechnical system of use," which is understood as "a system using combinations of hardware, people (and usually other elements) to accomplish tasks that humans cannot perform unaided by such systems" (Kline, 1985, 217).

As is usual in recent discussions on the Neutrality Thesis, I will primarily focus on technological artifacts. In Section 6, however, I will return to this topic and discuss the broader sociotechnical notion of technology as well. One advantage of focusing on technological artifacts is that participants can easily agree on some central examples: bridges, cars, hammers, sea dikes, etc. By having a shared set of examples to discuss, philosophers run less of a risk of talking past each other. However, beyond these examples, one finds surprisingly little about what exactly is meant by 'technological artifacts' in debates about the Neutrality Thesis.² In

² For example, Pitt consistently uses 'technological artifacts' in his 2014 article, but does not define this notion further. In other works, he uses a different notion of technology. In his 2000 book, *Thinking about Technology*, Pitt defines technology as "humanity at work". In more recent work, he argues that "technological infrastructures are those arrangements of people, structures and materials that make human activities possible" (2019, 64).

this light, it might be useful to draw some lessons from the recent philosophical literature on artifacts.

It is common to attribute two properties to artifacts: (1) they are the *intentional* products of human activity, and (2) have *functions* to do something (Hilpinen, 1992; Juvshik, 2021, 9313). Although artifacts are typically intentional products of human activity, some philosophers have argued that there are cases of unintentionally created artifacts (e.g. a path that has been unintentionally created by travellers) as well as artifacts that are not the product of human activities (e.g. a dam constructed by a beaver with the function of creating ponds) (Sperber, 2007, 125). Such cases are less relevant in the context of the debate on the Neutrality Thesis. There are two reasons for this. First of all, the kind of technologies that philosophers are primarily concerned with in the debate is *human* technology in specific. From this, it does not follow that it is uninteresting to think about non-human technology.³ It would, for example, be a necessary aspect of an evolutionary account of the development of human technology to look at the continuum between the products of both human and animal activity. It also makes sense to talk about technology that is not directly, but only indirectly, created by human beings. In this regard, one might consider cases in which AI systems develop their own technologies. Although it is good to be aware of these possibilities, focusing primarily on *human* technology gives the discussion a more disciplined focus and a clearly delineated subject matter. Where required, these restrictions can be relaxed in order to discuss more complicated cases. Secondly, focusing on intentionally created artifacts is often assumed in contemporary discussions on the Neutrality Thesis as well, given the focus on technological artifacts such as hammers, sea dikes, guns, etc. Here it is important to note that it is not because a technological artifact was intentionally created to perform a particular function that the artifact will always be used according to its intended function. There is an important difference between *intended* and *actual* function. This distinction will be more elaborately discussed in Section 5.

Having elucidated the notion of technological artifacts as the intentional products of human activity to perform a particular function, we can reformulate the Neutrality Thesis as follows:

(NT – TA) Technological artifacts are morally neutral.

A second step is to elucidate what it means for technological artifacts to be ‘morally neutral’. In his definition of the Neutrality Thesis, Pitt (2014, 90) talks about technological artifacts ‘not having’, ‘not having embedded in them’, or ‘not containing’ moral values. In order to understand these phrases, it is important to elucidate the notion of a moral value and the notion of what it means for technological artifacts to contain (or lack) moral values. For now, I will leave aside a discussion of the idea of what it means for technological artifacts to ‘contain’ or ‘have embedded in them’ moral values. The reason for this, as we will see, is that one main argument for The Neutrality Thesis depends on a specific interpretation

³ See Shew (2017) for a detailed study of (non-human) animal technology.

of such phrases by connecting talk of ‘containing moral values’ with the notion of empirical identifiability. This concept will be elaborately discussed in Section 5. However, it is important to already say something more about the concept of a value. What is a value?

In his discussion of the Neutrality Thesis, Pitt defines the notion of a value as “an endorsement of a preferred state of affairs by an individual or group of individuals that motivates our actions” (Pitt, 2014, 91). As a self-proclaimed pragmatist, Pitt proposes an *action-motivating* conception of values. In a recent criticism, Miller argues against this conception, stating that “the relation between values and motivations is not conceptually necessary” (Miller, 2021, 59). Against Pitt, Miller argues that one can adhere to a value by exhibiting a “*mere passive appreciation* without any motivation to act” (Miller, 2021, 59), and gives the example of the beauty of mathematics, which might be valued by someone who does not have any motivation to practice mathematics herself.

Although I agree with Miller that it makes sense to talk about values that are not action-motivating, one could respond by making a distinction between *thick* and *thin* values. A *thick* value would be a value that is, as a matter of fact, action-motivating. A *thin* value would be a value that is not action-motivating and which can be merely passively appreciated.⁴ Whether one would accept this response or not, I think that there is a more fundamental problem with Pitt’s definition. The problem is that he defines a value as an “*endorsement* of a preferred state of affairs” (Pitt, 2014, 91 - my italics). However, a value is not the endorsement itself, but rather the *object of (possible) endorsement*. We can ‘live in accordance with’, ‘promote’, or indeed ‘endorse’ a set of values. The notion of value that is at stake in the debate on the Neutrality Thesis is the one expressed by the (countable) *noun* ‘value’ (and not the *verb* ‘to value’), which is used to refer to *specific* values such as the values of autonomy, fairness, and loyalty. After all, discussions about the Neutrality Thesis concern the question of whether certain technological artifacts can be said to exemplify or go against certain moral values. These moral values are not the endorsements themselves, but the things or objects that can be endorsed. In other words, the relevant notion of values in discussions about the Neutrality Thesis is the notion of values as objects of possible endorsement.⁵

⁴ I leave open the possibility that there are values which can be thin or thick in different situations. For example, the aesthetic value of beauty can be action-motivating in certain situations (and thus amount to a *thick* value), whereas, in other situations, beauty can be merely passively appreciated (and thus amount to a *thin* value).

⁵ Two further remarks are necessary here. First of all, not all values are moral values. Non-moral values include for example aesthetic or prudential values. Given that the debate focuses on the question whether technology is *morally* neutral, the main focus in this paper is on *moral* values as well, even though it is sometimes useful to discuss examples of non-moral values when discussing the more general notion of ‘value’ as such. Secondly, I use a very lightweight conception of the notion of an ‘object’ here: objects can be but do not need to be *physical* objects. By calling values possible objects of endorsement, I do not wish to commit myself to the idea that values are physical objects. I do think this is in line with certain everyday uses of ‘objects’: for example, we might talk about mathematical objects without committing ourselves to the view that mathematical objects have physical properties (e.g. spatiotemporal location, shape, mass etc.).

A third and last point concerns the Neutrality Thesis' modal status. Most philosophers who argue against the Neutrality Thesis implicitly treat the Neutrality Thesis as a claim about the *necessary* connection between being a technological artifact and being morally neutral. Kroes and van de Poel (2014, 104) make this assumption explicit when arguing that the Neutrality thesis should be understood as the claim that technological artifacts *cannot* embody moral values. Similarly, Klenk has recently stated that “[d]efenders of the value-neutrality thesis [...] deny that the artefact itself *can* have value” (Klenk, 2021, 526 - my italics). Indeed, most arguments against the Neutrality Thesis proceed by giving examples of technological artifacts that seem to exemplify (or go against) moral values. Whether proponents of the Neutrality Thesis treat it in the same way is often hard to tell, given that a clearly developed argument in favour of the Neutrality Thesis is often missing. However, in his recent defence of the Neutrality Thesis, Pitt clearly suggests that he treats the Neutrality Thesis in this modally strong way when he claims that “the technologies themselves *cannot* in any legitimate sense embody values” (Pitt, 2014, 90 – my italics) and that “the case needs to be made for why values are the sorts of things artifacts cannot have in any *meaningful* way” (Pitt, 2014, 91).

With this in mind, we can reformulate the Neutrality Thesis as follows:

(NT – TA - N) Necessarily, technological artifacts do not embody moral values.

One nice consequence of making this modal assumption explicit is that the criteria for falsifying the Neutrality Thesis become clearer.⁶ Those who argue *against* the Neutrality Thesis do not have to be committed to the idea that there are *no* examples of ‘value-free’ technologies. Such critics might agree that certain examples to which proponents of the Neutrality Thesis refer are indeed morally neutral. All that they are committed to is that there are also possible cases in which technological artifacts are value-laden. In what follows, I will understand the Neutrality Thesis as a *neutrality thesis* about *technological artifacts* concerning the *necessary* connection between being a technological artifact and being morally neutral.

I will call this formulation the ‘strong version’ of the Neutrality Thesis. This version is central in most of the recent discussions of the Neutrality Thesis and, therefore, my primary focus will be on this formulation. This does not mean that the Neutrality Thesis cannot be understood in other ways. I have mentioned the fact that there are different possible notions of technology. ‘Broader’ formulations of the Neutrality Thesis would incorporate such broader notions. It is also possible

⁶ The relevant kind of necessity here is *metaphysical* necessity. Those who are suspicious of (heavy-weight) metaphysics can interpret the strong version of the Neutrality Thesis in accordance with Amie Thomasson’s recent modal normativist position (2020) concerning metaphysical necessity. According to Thomasson, metaphysically necessary claims should not be interpreted as stating ‘metaphysically deep’ facts about the world, but rather as disguised normative claims concerning the proper use of (for example) the predicates ‘is a technological artifact’ and ‘is morally neutral’. According to Thomasson’s account, the strong version of the Neutrality Thesis would convey the rule that one ought to be prepared to apply ‘is morally neutral’ to whatever particulars satisfying the predicate ‘is a technological artifact’. I remain neutral in this paper regarding the proper account of metaphysical modality.

to formulate what might be called ‘weaker versions’ of the Neutrality Thesis. One theoretical option would be to argue that technological artifacts can indeed embody values, but that they will always embody conflicting values. It could be argued that these values ‘cancel each other out’, and that therefore technological artifacts are morally neutral, despite the fact that they can indeed embody moral values. While it is unclear whether such a ‘calculus of values’ is intelligible (and, even if it would be intelligible, whether these values would *in every possible case* cancel one another out), this formulation amounts to a ‘weak version’ of the Neutrality Thesis. One could reject the ‘strong version’ but might still hold on to this weaker version of the Neutrality Thesis. In what follows, my focus will be on the strong version of the Neutrality Thesis. I do think, however, that the ‘narrow’ notion of technology as referring to technological artifacts is closer to the ‘broader’ sociotechnical notion of technology than is commonly thought. In Section 6, I will return to this exact issue.

3 Empirical Identifiability

Pitt’s argument in favour of the Neutrality Thesis can be summarized as follows (see Miller, 2021, 55):

- (P1) A necessary condition for technological artifacts to be morally non-neutral (and hence embody moral values) is that these moral values are empirically identifiable from technological artifacts;
- (P2) Moral values cannot be empirically identified from technological artifacts;
- (NT – TA - N) Necessarily, technological artifacts do not embody moral values.⁷

Miller correctly observes (2021, 58) that Pitt does not explicitly defend (P1). However, Miller claims that the truth of (P1) logically follows from Pitt’s definition of ‘value’. As we have seen, Pitt (2014, 91) defines a value as an endorsement of a preferred state of affairs by an individual or group of individuals that motivates action. If this is how ‘value’ is to be understood, Miller argues, Pitt “risks begging the question [...] by making values impossible to be embedded in material objects by definition” (Miller, 2021, 59). While I agree with this point, it does not undermine (P1). Instead, it seems to be directed at (P2); (P2) states that moral values cannot be empirically identified from technological artifacts. If values are endorsements, then they cannot be embedded in technological artifacts, and *hence* they cannot be empirically identified. What is question-begging is Pitt’s claim that moral values cannot be empirically identified from technological artifacts. And this claim is captured by (P2). In other words, I disagree that Miller’s point undermines (P1).

⁷ Note that Pitt’s original argument does not exclusively target the strong version of the Neutrality Thesis. Pitt also argues that *if* technological artifacts could embody moral values, then it would be difficult (or even impossible) to identify these values so that the claim that artifacts embody values would either be insignificant or trivial. Miller captures this nuance in Pitt’s argument by formulating the conclusion as a disjunctive claim. Given that my primary target is the strong version of the Neutrality Thesis, I have left out this second disjunct in my reformulation of Pitt’s original argument.

After all, it is possible to hold, as Pitt exactly does, that the empirical identifiability of values is necessary for technological artifacts to embody moral values (this is the first premise), even though one holds that it is impossible to empirically identify moral values from technological artifacts (this is the second premise).

Miller further nuances his point and claims that “[i]t might be argued that regardless of the conceptual relations between values and motivations, VNT1 [(P1) in my discussion] is still correct” (Miller, 2021, 59). He then goes on to give two counterexamples that are indeed genuinely directed at (P1). The first example is the following:

“For example, a blind person may value excellence in archery in a way that impacts her life, for example, she may collect memorabilia associated with great archers and admire them at bedtime, but she may be unable to recognize a good archer. Namely, an archer may embody the value of excellence while she cannot recognize it in the archer” (Miller, 2021, 59).

I do not think that this first example undermines (P1). First of all, (P1) concerns *empirical* identifiability and not *visual* identifiability. Whereas it is clear that a blind person cannot *visually* identify or recognize a good archer, whether or not she cannot *empirically* identify a good archer is far less clear. Sight is one important way of becoming empirically informed about our surroundings, but it is certainly not the only way. The blind person referred to in the example seems to be very passionate and knowledgeable about archery. Perhaps she even used to be an excellent archer herself. It definitely seems possible that such an admirer of archery could recognize a good archer in the way that the archer speaks about archery, by feeling the composure of the archer and the relaxedness and steadiness with which she draws her bow, and even by the time needed to shoot the arrow.

Secondly, even if there would be cases in which a particular person could not empirically identify a particular value, it would still not necessarily follow that (P1) is false. Empirical identifiability is a notion that lacks concreteness without some sort of specification as regards the question to *whom* something must be empirically identifiable. It would be impossible to empirically identify, from the lines of code that represent Twitter’s (X’s) algorithms, whether or not Twitter (or X) artificially boosts its owner’s tweets for someone who does not have any programming skills. From this, though, it would not follow that the fact that Twitter (or X) artificially boost its owner’s tweets is a fact that is empirically *unidentifiable* in the relevant sense. The empirical identifiability that is at stake could be a kind of empirical identifiability *for people with sufficient programming skills*. While there is some leeway in interpreting the relevant notion of ‘empirical identifiability’ in (P1), it would be extremely uncharitable to interpret a defender of (P1) as being committed to the idea that moral values have to be empirically identifiable by *anyone*. Even if there are cases where a value is empirically unidentifiable by a particular person, this would not necessarily count as a counterexample to (P1).

This point is also relevant to Miller’s second argument against (P1). Miller makes the interesting point that “values in technology are so effective *because* they are often hardly empirically recognizable” (Miller, 2021, 59). I agree with this claim but, again, it does not undermine (P1). (P1) states that a necessary condition for

technological artifacts to be value-laden is that these values be empirically identifiable. This is compatible with the claim that there are technological artifacts that are value-laden, but where empirically identifying these values is possible but difficult. In order to be a genuine counterexample to (P1), one needs to give an example of an instance in which a technological artifact is value-laden and where the relevant moral value is empirically unidentifiable. Miller does not provide such an example. Therefore, he fails to effectively argue against (P1).

This does not mean that (P1) is without its problems. In order to fully evaluate (P1), more would need to be said about the precise meaning of empirical identifiability. Which kind of identifiability is at stake? Should moral values be identifiable by adult human beings who have their rational faculties more or less intact? Is the notion of identifiability relative to being identifiable by experts in the relevant, technological field? Pitt does not give any answer to these questions. My criticism of Miller's discussion of (P1) should, therefore, not be read as a defence of (P1). My point is that if one wishes to reject Pitt's argument for the Neutrality Thesis, then these are the wrong reasons to do so. The best way to reject the conclusion of the argument is to deny (P2), as I will now proceed to argue.

4 Technological Artifacts Are Not Just Physical Objects

(P2) states that moral values cannot be empirically identified from technological artifacts. Pitt justifies (P2) as follows:

"[...] if we look at the actual physical thing – the roads and bridges, etc. where are the values? I see bricks and stones and pavement, etc. But where are the values – do they have colors? How much do they weigh? How tall are they or how skinny?" (Pitt, 2014, 95)

While this is not a clear-cut argument for (P2), it does express an important intuition which underlies many endorsements of the Neutrality Thesis. In the end, such philosophers argue that technological artifacts, such as roads and bridges, are nothing more than physical objects. Physical objects can be spatio-temporally located; they have a mass and a particular shape. Values cannot be embedded in physical objects because values are human constructions. Physical objects are indifferent towards human values. The same physical object might be used to serve different values, and they might mean different things to different people (Pitt, 2014, 94). Hence, technological artifacts, which are nothing but physical objects, are morally neutral. They do not embody any moral values.

Miller has recently countered this argument. He argues that "sometimes values are directly readable off design documents or material artifacts" and that "[t]echnology may also have expressive meaning that *implicitly* conveys values" (Miller, 2021, 61). However, the obvious response for the defender of the Neutrality Thesis is to state that this does not count as an argument against the neutrality of technological artifacts. After all, there is a difference between the physical object itself and the content that is 'implicitly conveyed'. The physical object is indifferent

towards the content conveyed: the content is only there because it is treated as such by human interpreters.

Miller considers such a response, but argues that “[t]his objection, however, wrongly assumes that content and the material means that stores, processes, or delivers it are sharply separable from each other. But content cannot be expressed without material means [...]” (Miller, 2021, 61). However, this is not a convincing response to a proponent of the Neutrality Thesis. A defender of the Neutrality Thesis can agree that there cannot be content without the material means to express content, but simply argues that the material itself does not express content, and therefore does not embody any values. It might be argued that it still remains the case that no values are embodied, given that technological artifacts are nothing more than physical objects.

Miller’s last argument is not just that content cannot be expressed without material means, but that certain contents, which express certain values, can *only* be expressed by *specific* material means. For example, “[o]nly a danger sign with certain physical properties embeds the value of safety. A flashing sign that distracts drivers from the danger from which it is supposed to warn them or an unreadable sign does not embed safety” (Miller, 2021, 61). Again, Miller’s observation is correct, but it does not undermine the position of a proponent of the Neutrality Thesis. A defender of the Neutrality Thesis can agree that certain value-laden messages can only be expressed by making use of particular kinds of material means, but that this does not change the fact that the material means themselves are indifferent towards moral values.

The fundamental problem is a different one, and it concerns the nature of technological artifacts. What underlies Pitt’s argument against the empirical identifiability of moral values from technological artifacts is that technological artifacts are to be understood as *mere* physical objects. The properties of such physical objects can be characterized by describing their spatio-temporal properties, their mass, and their shape. The proper vocabulary to describe these features is the vocabulary of theoretical physics. Values are not part of the physical world in this sense; the concept of values does not feature in the vocabulary of theoretical physics. Therefore, it does not make sense to say that values are ‘empirically identifiable’ from, or are ‘locatable’ in, technological artifacts understood as mere physical objects.

I dispute the assumption that technological artifacts can be fully described using a physicalistic vocabulary only. While it is true that physical descriptions are necessary to properly describe technological artifacts, they are not sufficient. Technological artifacts also have functional properties. In other words, technological artifacts have a hybrid or ‘dual nature’: “they are (i) designed physical structures, which realize (ii) functions, which refer to human intentionality. This conceptualisation of technical artefacts, as physical and as functional objects, *combines* two fundamentally different ways of viewing our world” (Kroes & Meijers, 2006, 2).

Kroes and Meijers occasionally also talk about ‘physical’ and ‘intentional’ conceptualizations. While there is nothing wrong with this as such, I think that it is theoretically useful to clearly emphasize the distinction between functions and intentions as well. The main reason for this, as I will argue in Section 5, is that technological artifacts can have *unintended* functions that are morally relevant. Kroes, for example,

distinguishes between (1) the structural (or physical) properties, (2) the functional properties of a technological artifact, and adds that these properties are embedded in a (3) context of intentional human action (Kroes, 2010, 56–57). This tripartite structure is useful for my purposes because it more clearly separates technological artifacts' functional and intentional components. While artifacts are distinguished from mere physical objects by possessing (originally) intended functions, this does not preclude the possibility of them having non-intended functions as well.

The main argument against a physicalist reduction of technological artifacts is that there is a 'logical gap' (Kroes, 2006, 141) between the physical and functional descriptions of a technological artifact. An artifact can function *properly* or *improperly*, and is thus subject to a kind of *normative assessment* (Franssen, 2006). A hammer or car can *malfunction*: it makes sense to distinguish between 'good' and 'bad' hammers or cars. In so far as normative descriptions cannot be reduced to physical descriptions, functional descriptions are necessary to capturing this aspect of technological artifacts.

This point is further supported by the claim that functional descriptions are omnipresent in actual engineering practices (Kroes, 2010). Because engineers are "experts in designing, making, analysing and describing technical artefacts" (Kroes, 2010, 52), their descriptions of artifacts deserve to play an important evidential role in a philosophical account of the nature of technological artifacts. While the 'logical gap' argument perhaps does not count as a knock-down argument for a reductive physicalist, I do think that it shifts the burden of proof onto the physicalist to come up with a revisionary approach to bridge the 'logical gap' between functional and physical properties. Furthermore, a *liberal* or *soft* naturalist (rather than a *hard* or *reductive* naturalist) can point to the many established scientific practices in which the empirical identification of functions plays an essential role.⁸ Taking these practices seriously lends further support to the idea that functional properties deserve to be included in the naturalist's worldview.

5 The Empirical Identifiability of Functions and Values

I have argued that Pitt's assumption underlying (P2), the assumption that technological artifacts are nothing more than physical objects, is highly disputable. I have argued against this by claiming that technological artifacts are not merely physical objects, but are instead the intentionally created products of human activity with a particular function to do something. Technological artifacts cannot be described using a merely physicalistic vocabulary: the concepts of *intention* and *function* are also necessary to describe technological artifacts properly. In this section, I aim to achieve three things. First, I will distinguish between different types of functions that technological artifacts can have (5.1). Secondly, I will argue about the sense in which these different types of artifact functions can be identified empirically (5.2).

⁸ The distinction between strict or reductive and liberal or soft naturalism is due to P.F. Strawson (1985, 5).

Thirdly, I will argue that the empirical identifiability of artifact functions leads to a natural way to defend the empirical identifiability of moral values from technological artifacts. This leads to a rejection of Pitt's second premise (5.3).

5.1 Four Types of Functions

There are two central distinctions in the philosophical literature on the concept of function that are especially relevant for my purposes here.⁹ The first distinction is the one between (1) system functions and (2) aetiological functions. System functions are contributions that are made by a particular entity to the capacities of a larger system of which it is a part (Cummins, 1975). For example, a car battery's system function is to contribute to a car's larger capacity to transport people, just like how the screws in a bookshelf contribute to the bookshelf's larger capacity to support and display books. Most entities with system functions have more than one system function: the car battery not only contributes to the car's capacities, but also presumably to the car owner's capacities, her family, and the company for which she might work. The notion of a system function is limited to the actual or current dispositions or capacities of a system and does not depend on its causal history. The explanatory value of ascribing a system function to an entity x is that such an ascription explains the capacities of larger systems of which x is a part. It does not explain the existence of x .

Aetiological accounts (Millikan, 1984; Wright, 1973) differ in this respect. One classic example of an aetiological function in the biological realm is the function of the heart to pump blood in order to transport nutrients to, and waste away from, cells. The function of pumping blood is an effect of a natural selection mechanism that causally explains the existence of hearts in the first place. The explanatory value of ascribing an aetiological function to an entity x is, thus, that such an ascription explains that entity's existence. Artifacts can also have aetiological functions. A light switch's function to turn on and off lights causally explains the existence of light switches in the sense that its function is an effect not of a natural selection mechanism, but rather of intentional action. Artifact functions differ from biological functions in the sense that the former typically are the result of (design) intentions, while the latter are not.

⁹ A distinction is made between intentionalist theories, aetiological (or selectionist) theories, and system function theories in the literature (for an overview of the debate, see Preston, 2009; Artiga, 2023, 1537–1539). All of these theories face counterexamples. Intentionalist theories have difficulties accounting for the unintended functions of artifacts. Aetiological theories have problems accounting for phantom functions and for the fact that the first exemplar of a new artifact can have a function. System functions theories are often considered to be too inclusive (and thus incapable of distinguishing between 'proper' and 'accidental' functions) and as being incapable of allowing there to be dysfunctions. My goal here is not to evaluate these theories or to defend a general theory of artifact function. Instead, the aim is to apply a number of important conceptual distinctions in order to bring into view different types of functions that artifacts can possess that are empirically identifiable and which can play a potential role in the value-ladenness of technological artifacts. While my sympathy lies with pluralist (see Preston, 1998 for a classic account) and hybrid theories (Artiga, 2023), a full defence of these approaches is well beyond the scope of this paper.

As Preston has convincingly argued, the notions of system function and aetiological function “are not competitors, but are complementary” (Preston, 1998, 217). In some cases, a system function of an entity can also be an aetiological function. For example, the heart’s function to pump blood is both a system function (as it contributes to the capacity of the blood circulatory system) as well as an aetiological function (as it causally explains the existence of hearts). Similarly, it is plausible that the car battery’s system function to contribute to the larger capacities of a car is also its aetiological function in the sense that this function plays an explanatory role with regard to the existence of car batteries in the first place.

The second distinction is the one between (1) intended function and (2) non-intended function. ‘Intended function’ can refer either to the intentions of the original designer (inventor) or the intentions of an (occasional) user of the artifact. I will mark this distinction by distinguishing between an artifact’s ‘originally intended function’ and the ‘intended use function’. These two types of intended functions can play different explanatory roles with regard to the existence of artifacts. After all, the notion of ‘existence’ can either refer to an entity’s *emergence* or to its *maintenance* (its continued existence) (Kitcher, 1993; Köhler, 2022). Whereas the ‘originally intended function’ explains the *emergence* of an artifact, an ‘intended use function’ can play a significant role in the explanation of an artifact’s continued existence.¹⁰ These two subtypes of ‘intended function’ are to be distinguished from ‘non-intended functions’. As I have argued, biological functions are typically non-intended aetiological functions that explain the existence of the relevant biological trait or entity. However, as I will argue below, there is also an important sense in which artifacts can be said to have non-intended aetiological functions.

If we combine these two distinctions, there are four possible combinations which lead to four types of functions that artifacts can be said to have: (1) intended aetiological functions, (2) non-intended aetiological functions, (3) intended system functions, and (4) non-intended system functions.

Consider first the class of intended aetiological functions. While the natural selection mechanism is a non-intentional mechanism, the emergence of artifacts is typically explained by an intentional mechanism. As we have seen, ‘intended function’ can refer both to an artifact’s ‘originally intended function’ and to its ‘intended use function’. The ‘originally intended function’ of sea dikes in the Netherlands is to protect people from floodings, and the originally intended function of city benches with anti-homeless spikes is to prevent homeless people from sleeping on them. These aetiological functions explain the emergence of sea dikes and city benches with anti-homeless spikes, respectively.

Originally intended functions can diverge from intended use functions. As Preston (2009, 215) has argued, one central feature of artifacts is that they are *multiply utilizable*. This means that artifacts can be used to perform different functions, including functions that were not originally intended. One clear example is dynamite, which has been intentionally used for the purposes of killing people, even

¹⁰ Of course, in cases where the ‘originally intended function’ and the ‘intended use function’ coincide, they can explain both the relevant artifact’s *emergence* and its *maintenance*.

though most accounts of Alfred Nobel's original intentions deny that he intended it to be used for such purposes. The function of killing people is, thus, an intended use aetiological function in so far as this function plays a significant role in the explanation of dynamite's continued existence. To sum up, there are two subtypes of intended aetiological functions: (a) originally intended aetiological functions and (b) intended use aetiological functions.

The second type of function is that of non-intended aetiological functions. This interesting class of functions includes what the sociologist Robert Merton (1968) called 'latent functions'. These are functions that are neither intended nor recognized by a community's members. One example is the function of a totem pole to reinforce group identity in a tribe, even if members of the tribe are unaware of this function and did not specifically intend the totem to have this function (Artiga, 2023, 1537). Such unawareness is compatible with the claim that the function to reinforce group identity plays a significant role in the continued existence of totem poles. In general, artifact functions that promote a group's social cohesion are good examples of non-intended aetiological functions. Certain religious objects, but also football or basketball shirts, can have social functions to strengthen group ties, even though these functions need not be consciously intended by anyone and can remain unrecognized by all members of a community.

The third and fourth types of functions concern system functions rather than aetiological functions. Given that system functions can also be aetiological functions, some of the abovementioned examples also count as examples of system functions. Sea dikes in the Netherlands contribute to the capacities of a system of flood control in conjunction with drainage ditches, pumping stations, and canals. City benches with anti-homeless spikes contribute to a larger system of excluding the powerless from places in which the powerful do not want them. Gas chambers contribute to the system, grounded in the Nazi ideology, whose main aim was to exterminate Jewish people. The functions of these artifacts contribute to a larger system that also comprises many other technological artifacts. Just as with aetiological system functions, these functions can be both intended or not intended.

The notion of a system function is extremely inclusive because system functions do not need to have any significant explanatory value with respect to the artifact's existence. This is especially the case for non-intended system functions. Most artifacts will play a marginal role in terms of contributing to the capacities of a huge number of systems. The water bottle on someone's desk contributes not only to the capacities of the person who drinks from it to stay hydrated, but presumably also to the capacities of the owner's family, sports club, company, the global economy, etc. For my purposes, the notion of a system function plays one specific explanatory role in this paper. After all, one important class of artifacts that has been said to embody moral values are artifacts that contribute to a larger system of harmful social inequalities. Importantly, such system functions need not play any significant role in causally explaining the existence of the artifact. Miller, for example, discusses the 2009 controversy about the HP camera which only "tracks the movement of a white woman's face but not a black man's face" (Miller, 2021, 61). In recent years, there has been an increased awareness of technologies that reflect and perpetuate social inequalities: smartphones, offices, and cars that are designed for

men and harm women (Criado Perez, 2019), facial technology that fails to recognize darker-skinned women (Najibi, 2020), or technologies that reflect and sustain racism more generally (Benjamin, 2019). Even when such contributions to harmful social structures happen *unintentionally*, it can be argued that these are nevertheless examples of immoral technologies. In what follows, I will limit my focus on the class of non-intended system functions that are not necessarily also aetiological functions.

5.2 The Empirical Identifiability of Functions

Functions are empirically identifiable. Empirical research in evolutionary biology constitutes a clear example in the case of non-intended aetiological functions in the biological realm. The dozens of samples of birds and other animals that Darwin collected during his famous 5-year voyage on his ship, the *Beagle*, to the Galápagos Islands, would constitute the empirical backbone of his theory of evolution. His observation of different species of finches on the Galápagos islands was essential to explaining the existence of adaptations (i.e. features of an organism which are functionally designed through the process of natural selection). The denial of the empirical identifiability of such adaptations would amount to a denial of evolutionary biology's scientific legitimacy.

There is no reason to think that empirical identifiability is solely limited to biological functions. Artifact functions can also be empirically identified. In the case of intended aetiological functions, empirical identifiability is made possible for example by historical studies. One nice example of such research concerns the underlying intentions of urban planner Robert Moses' order to build low bridges in Long Island. On the basis of "evidence provided by Robert A. Caro in his biography of Moses" (Winner, 1980, 123), Winner famously argued that these bridges reflected Moses' racial prejudices. Other writers have disputed the historical accuracy of this claim, however. Bernward Joerges (1999, 416–419) has disputed the evidence brought forward by Caro and referenced by Winner. This is what good history research is all about: to research whether historiographical criteria are satisfied, which includes an investigation of the accuracy of sources and possible alternative explanations. On the basis of such qualitative, empirical research, historians gather evidence in support of (or against) a particular claim concerning the intended functions of technological artifacts. Although empirically identifying intentions in such a way is often a difficult matter (and demands proper training), denying the empirical identifiability of such intended functions amounts to a denial of the legitimacy of a well-established scientific practice.

Non-intended functions can also be empirically identified. As Artiga has argued, the recognition of such functions "is fundamental for anthropology and sociology, since finding out that an object or behavior lacks a manifest function should not stop research" (Artiga, 2023, 1537). Sociological approaches that study the non-intended aetiological functions of religious objects or objects to commemorate ancestors can reveal hidden social functions that not only might never have been intended, but which also need not be transparent to a particular community's members. Even though the results of sociological and anthropological research are not

infallible (empirical sciences are simply not infallible), examples of such kinds of research show that there are scientifically respectable ways to empirically identify non-intended aetiological functions. A denial of the empirical identifiability of such functions would also amount to the denial of the legitimacy of some well-established scientific practices. Furthermore, it is also possible to empirically identify the harmful non-intended system functions of HP cameras, smartphones, or cars to contribute to systems of social inequalities. Such research ought to be done on the basis of a systematic and transparent gathering of relevant data, the proper use of statistical methods, a critical assessment of alternative explanations, and a general sensitivity to, and awareness of, persistent social and economic inequalities.

In other words, there are many, non-mysterious ways in which (both intended and non-intended) functions are empirically identifiable. Indeed, the possibility of empirically identifying different types of functions is central to the scientific methodology of many different disciplines. From the fact that empirical identification of functions is sometimes difficult, it does not follow in any way that it is impossible. Denying the empirical identifiability of functions would, thus, amount to a denial of some well-established scientific practices. I, therefore, take myself to have presented sufficient evidence in support of the claim that functions are, as a matter of fact, empirically identifiable.

5.3 The Empirical Identifiability of Moral Values

If functions are empirically identifiable, then a natural way to defend the empirical identifiability of moral values becomes available. One way of seeing this is through the concept of *intention*. Many of the classic examples of technological artifacts discussed in the literature embody moral values because these are cases where there is little to no doubt about the originally intended function and most of their uses are in line with this originally intended function. Therefore, I will first focus on artifacts with intended (aetiological) functions. Afterwards, I will consider technologies with non-intended system functions and suggest that even in the absence of clear intentions, it might nevertheless be defensible to hold that these technologies can embody moral values as well.

The central class of cases of technological artifacts embodying moral values concerns artifacts with intended aetiological functions. After all, intentions can be the objects of moral (dis)approval. The intention to create artifacts with a particular kind of function assumes some agential *control* over the artifact. Empirically identifying an intended aetiological function can then expose the existence of the artifact as causally explained by, for example, the creator's racist ideology (or group of creators) or by a morally praiseworthy intention to protect a moral value. It is no accident that the standard examples given in the literature (the gas chambers during the Holocaust, city benches with anti-homeless spikes, sea dikes etc.) are cases where there is little doubt about the artifact's originally intended function. Moral values can be empirically identified because the originally intended functions of artifacts are empirically identifiable. But not only the original intention plays a role in ascribing moral values to

technological artifacts. Most actual intended use functions of the abovementioned classic examples are in line with the originally intended function as well: the gas chambers during the Holocaust have effectively been used mainly for the realization of its racist goals, sea dikes are mainly used to protect people from floodings, and city benches with anti-homeless spikes predominantly prevent homeless people from sleeping on them. When these empirical truths are added to the empirical truths about the originally intended functions of these artifacts, there is sufficient empirical evidence to justify the claim that technological artifacts can indeed embody moral values.

The distinction between ‘originally intended function’ and ‘intended use function’ allows for a response to an objection raised by Klenk (2021, 529–533) against a similar approach developed by Kroes and van de Poel (2014). Van de Poel and Kroes also support their claim that artifacts can embody moral values by referring to the intended functions of artifacts. But, as Klenk argues, they limit their notion of intended functions to what is intended by the designer of the artifact. In other words, their account of intended function is limited to what I have described as ‘originally intended function’. Because of this limitation, their intentional history account is unable to account for changes in embedded values. As I have argued in the previous section, however, my account allows for both types of functions. If the originally intended function of an artifact is something morally blameworthy (praiseworthy) but most of its subsequent uses are morally praiseworthy (blameworthy), it would be plausible to maintain that an artifact embodies a moral value which is different from the moral value that was initially embodied by the artifact.

What about non-intended functions? It is sometimes said that certain technologies are immoral because these technologies reflect and further perpetuate social inequalities. What is interesting about these types of cases is that such claims can be made even if the intentions of the creators of these technologies were not morally bad. However, if these technologies do not embody moral values because of their *intended* functions, then in what sense can it be said that these technologies embody moral values at all?

One straightforward answer is that they go against the moral value of *fairness* because these artifacts contribute to social inequalities. After all, it is social *inequality* that is being perpetuated. There is no doubt that the products of human design can contribute to social inequalities, despite the lack of bad intentions or even in the presence of good ones. While there are indeed technologies “that explicitly work to amplify hierarchies,” there are many “that ignore and thus replicate social divisions, and a number that aim to fix racial bias but end up doing the opposite” (Benjamin, 2019, 8). Non-intended, but harmful, system functions that replicate social inequality can be, and usually have been (cf. Benjamin, 2019; Criado Perez, 2019; O’Neil, 2011), attributed to particular kinds of technologies on the basis of empirical evidence. Lastly, note that the lack of a clear intention does not entail a lack of responsibility. Although the designers of technologies that unintentionally contribute to social inequality might have been ignorant, there are cases for which ignorance is *culpable*. Where technologies in specific are concerned, which are new and to be used by many people, those who create these technologies have a huge responsibility to assess the potential harmful consequences of their creations.

I conclude from this that moral values are empirically identifiable from technological artifacts. The more traditional examples in the literature are cases for which there is little doubt about the moral or immoral aims of technological artifacts' originally intended functions. Furthermore, these examples are all cases for which most of the intended use functions are in line with the originally intended function. Given that both types of intended functions can be empirically identified, I have argued that moral values embodied by technological artifacts are also empirically identifiable. Alongside those 'more traditional' examples, I have also discussed the interesting class of artifacts that do not embody moral values because of their intended functions, but because of their harmful *unintended* system functions that contribute to patterns of social inequality. Even though such research is often hard, the difficulty that this kind of research incurs does not undermine the possibility that, when everything goes well, these harmful system functions can also be empirically identified.

Therefore, (P2) is false. To sum up, the following three claims, for which I have argued in this section and in Section 4, support my rejection of (P2):

- (1) Technological artifacts are the products of intentional design with particular functions (intended and non-intended functions) to perform certain tasks;
- (2) Functions (both intended and non-intended functions) are empirically identifiable;
- (3) If functions are empirically identifiable, then moral values can be identified from technological artifacts.

From (1)-(3) it follows that:

(not-P2) Moral values can be empirically identified from technological artifacts.

If (P2) is false, then it also follows that Pitt's argument for the Neutrality Thesis is unsound.

6 Technology and Responsibility

Not only does my account undermine one prominent argument for the Neutrality Thesis, it also offers strong support for technological artifacts' possible value-ladenness in general. A clear path towards constructing and legitimizing counterexamples against the Neutrality Thesis opens up by arguing that technological artifacts cannot be merely reduced to physical objects. Technological artifacts have originally intended functions and intended use functions that can be empirically identified. Given that intentions are possible objects of moral (dis)approval, technological artifacts can be said to embody moral values if their originally intended functions are morally bad (or good) and most of its intended use functions deserve the same moral evaluation. While these cases of *intended* functions were central to my argument, I also showed the sense in which moral values can be empirically identified on the basis of the *unintended* functions of technological artifacts. Certain technologies can contribute to social inequalities even in the absence of bad intentions or in the presence of good ones.

One final point concerns the concept of technology. Debates on the Neutrality Thesis tend to focus on ‘technological artifacts’. The examples discussed range from sea dikes and bridges to city benches with anti-homeless spikes and gas chambers. Even though I agree that the focus on technological artifacts is a legitimate use of ‘technology’, it is definitely a *narrow* use of this term. As I made clear in Section 2, there are broader notions of technology as well. ‘Technology’ might refer not just to individual artefacts, but also to a network of artifacts in combination with computer programs, organizational structures, and the people that make up these structures and execute these computer programs. Kline (1985) calls this *broader* notion of technology the ‘sociotechnical’ one, and one of this term’s merits is that it highlights the deep intertwining between technology’s technical and social aspects. Alternatively, as Leo Marx argues, a “prominent feature of these complex, ad hoc systems is the blurring of the borderlines between their constituent elements – notably the boundary separating the artifactual equipment (the machinery or hardware) and all the rest: the reservoir of technical – scientific – knowledge; the specially trained workforce; the financial apparatus; and the means of acquiring raw materials” (Marx, 2010, 568).

There is certainly a sense in which the distinction between technological artifacts as discrete entities and the larger systems in which they operate (often systems of human relationships) is artificial. Technological artifacts usually have (intended and non-intended) system functions: they contribute to the capacities of systems that reach far beyond them. My suspicion is that one of the reasons why the debate has focused on the *narrow* use of ‘technology’ (understood as technological artifacts) is not just that examples of technological artifacts are easier to find and are more concrete, but also that the Neutrality Thesis is much harder to defend when a *broader* use of ‘technology’ is adopted. After all, it is much easier to argue that technology embodies moral values if ‘technology’ is immediately understood as not only referring to technological artifacts, but also to a broader network which includes the human beings who have developed and use this technology. It is, in the first place, human beings that defend and fight for the values they hold dear.

It is human beings that are the proper targets of our responsibility attributions as well. In the end, the moral values embodied by technological artifacts derive from human intentional action or from the culpable lack of awareness or sensitivity for social and economic inequalities. Of course, this does not exclude the possibility that there might, one day, be human creations which should then be properly recognized as fully-fledged members of our moral community. However, nothing about the claim that technology can be value-laden should be seen as a way to avoid responsibility by attributing it to some kind of non-human agency. On the contrary, one of the main functions of the claim that technological artifacts can embody moral values is to serve as an important reminder that technological artifacts expressly differ from the mere physical objects ‘already out there’ in the sense that they are the things for which their human creators should bear responsibility.

7 Conclusion

In this paper, I have argued that the Neutrality Thesis is incorrect. First of all, I clarified what the Neutrality Thesis is. Secondly, I analysed Pitt's recent argument in favour of the Neutrality Thesis. While I agree with Miller that this argument should be rejected, I argued that Miller rejects the argument for the wrong reasons. I argued that Pitt wrongfully reduces technological artifacts to physical objects and argued that the concepts of function and intention (which do not feature in a physicalistic vocabulary traditionally understood) are necessary to properly capture the concept of a technological artifact. I then distinguished between different types of functions that technological artifacts can be said to possess, and argued that these types of functions are empirically identifiable. This richer and more accurate understanding of technological artifacts not only led to a rejection of Pitt's argument's second premise, but also gave independent support for the thesis that technological artifacts can be value-laden.

Acknowledgements I would like to thank Fernando Braccacini, Stefaan Cuypers, Benjamin De Mesel, Jozefien Gielen, Sanne Huysmans, Iro Kyriakidou, Lode Lauwaert and two anonymous reviewers of this journal for their excellent comments and helpful feedback.

Authors' Contributions Sybren Heyndels (SH) is the only author of this paper

Funding Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek), Award number: 1206321N.

Data Availability N/A.

Declarations

Ethics approval and consent to participate N/A

Consent for publication N/A.

Competing interests The author has no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Artiga, M. (2023). A Dual-Aspect Theory of Artifact Function. *Erkenntnis*, 88, 1533–1554.
- Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity.

- Criado Perez, C. (2019). *Invisible Women: Exposing Data Bias in a World Designed for Men*. Chatto & Windus.
- Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, 72, 741–765.
- Franssen, M. (2006). The Normativity of Artefacts. *Studies in History and Philosophy of Science Part A*, 37(1), 42–57.
- Fry, H. (2018). *Hello World: How to Be Human in the Age of the Machine*. Random House.
- Hilpinen, R. (1992). On artifacts and works of art. *Theoria*, 58(1), 58–82.
- Joerges, B. (1999). Do Politics Have Artefacts? *Social Studies of Science*, 29(3), 411–431.
- Juvshik, T. (2021). Artifacts and mind-dependence. *Synthese*, 199(3–4), 9313–9336.
- Katz, E. (2005). On the neutrality of technology: The Holocaust death camps as a counter-example. *Journal of Genocide Research*, 7(3), 409–421.
- Kitcher, P. (1993). Function and Design. *Midwest Studies in Philosophy*, 18, 379–397.
- Klenk, M. (2021). How Do Technological Artefacts Embody Moral Values? *Philosophy & Technology*, 34, 525–544.
- Kline, S. J. (1985). What is technology? *Bulletin of Science, Technology & Society*, 5(3), 215–218.
- Köhler, S. (2022). What is (Neo-) Pragmatists' Function? *Australasian Journal of Philosophy*, 1–17.
- Kroes, P. (2006). Coherence of structural and functional descriptions of technical artefacts. *Studies in History and Philosophy of Science Part A*, 37(1), 137–151.
- Kroes, P. A. (2010). Engineering and the Dual Nature of Technical Artefacts. *Cambridge Journal of Economics*, 34, 51–62.
- Kroes, P. A., & Meijers, A. (2006). Introduction: The Dual Nature of Technical Artefacts. *Studies in History and Philosophy of Science*, 37, 1–4.
- Kroes, P. A., & van de Poel, I. R. (2014). Can technology embody values? In P. Kroes & P. P. Verbeek (Eds.), *The moral status of technical artefacts* (pp. 103–124). Springer Netherlands.
- Marx, L. (2010). Technology: The Emergence of a Hazardous Concept. *Technology and Culture*, 51(3), 561–577.
- Merton, K. (1968). *Social Theory and Social Structure*. The Free Press.
- Miller, B. (2021). Is technology value-neutral? *Science, Technology, & Human Values*, 46(1), 53–80.
- Millikan, R. G. (1984). *Language, thought, and other biological categories*. The MIT Press.
- Najjibi, A. (2020). *Racial Discrimination in Face Recognition Technology*. Retrieved 18 August, 2023: <https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/>
- O'Neil, C. (2011). *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. Penguin Books Ltd.
- Pitt, J. C. (2000). *Thinking About Technology: Foundations of the Philosophy of Technology*. Seven Bridges Press.
- Pitt, J. C. (2014). “Guns don’t kill, people kill”; Values in and/or around technologies. In P. Kroes & P. P. Verbeek (Eds.), *The moral status of technical artefacts* (pp. 89–101). Springer Netherlands.
- Pitt, J. C. (2019). *Heraclitus Redux: Technological Infrastructures and Scientific Change*. Rowman & Littlefield.
- Preston, B. (1998). Why is a Wing Like a Spoon? A Pluralist Theory of Function. *The Journal of Philosophy*, 95(5), 215–254.
- Preston, B. (2009). Philosophical Theories of Artifact Function. In: A. Meijers, *Philosophy of Technology and Engineering Sciences (Handbook of the Philosophy of Science, Volume 9)*. Amsterdam: Elsevier.
- Shew, A. (2017). *Animal Constructions and Technological Knowledge*. Lexington Books.
- Sperber, D. (2007). Seedless grapes: Nature and culture. *Creations of the mind: Theories of artifacts and their representation* (pp. 124–137). Oxford University Press.
- Strawson, P. F. (1985). *Skepticism & Naturalism. Some Varieties*. Routledge.
- Thomasson, A. (2020). *Norms and Necessity*. Oxford University Press.
- Veletsianos, G. (2014). *On Noam Chomsky and technology's neutrality*. Veletsiaonis.com. Retrieved 20 April, 2023, from <https://www.veletsianos.com/2014/01/23/on-noam-chomsky-and-technologys-neutrality>
- Winner, L. (1980). Do artifacts have politics?. *Daedalus*, 121–136.
- Woolgar, S., & Cooper, G. (1999). Do Artefacts Have Ambivalence? Moses' Bridges, Winner's Bridges and Other Urban Legends in S&TS. *Social Studies of Science*, 29(3), 433–449.
- Wright, L. (1973). Functions. *The Philosophical Review*, 82(2), 139–168.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.