Génova, G., Quintanilla Navarro, I. Discovering the Principle of Finality in Computational Machines. Found Sci 23, 779–794 (2018)

DOI: https://doi.org/10.1007/s10699-018-9552-4

Foundations of Science 23(4):779–794, December 2018.
(Published online: 13 February 2018)
The final publication is available at Springer via:
http://dx.doi.org/10.1007/s10699-018-9552-4

Discovering the principle of finality in computational machines

Gonzalo Génova
Departamento de Informática, Universidad Carlos III de Madrid
Avda. Universidad 30, 28911 Leganés (Madrid), Spain
ggenova@inf.uc3m.es / +34 91 624 8846
ORCID 0000-0003-0299-286X

Ignacio Quintanilla Navarro
Departamento de Teoría e Historia de la Educación, Universidad Complutense de Madrid
Avda. Rector Royo Vilanova s/n, 28040 Madrid, Spain
ignacioq@ucm.es

Abstract: In this essay we argue that the notion of machine necessarily includes its being designed for a purpose. Therefore, being a mechanical system is not enough for being a machine. Since the experimental scientific method excludes any consideration of finality on methodological grounds, it is then also insufficient to fully understand what machines are. Instead in order to understand a machine it is first required to understand its purpose, along with its structure, in clear parallel with Aristotle's final and formal causes. Obviously, purpose and structure are not machine components that can physically interact with other components; nonetheless they are essential to understanding their operation. This casts an interesting light on the relationship between mind and body: for just as an artifact's finality and structure explain its operation, so also consciousness is the explanation —not the efficient cause— of specifically human behavior. What machines and human beings have in common is that, in order to understand them, it is necessary to appeal to the principle of finality. Yet while finality is given and extrinsic in the case of machines, we human beings are characterized by the ability to self-propose our own ends. Since the principle of finality is essential to understanding the production of machines, the traditional view in modern Western philosophy that finality lies beyond the scope of objective/scientific knowledge should be rectified to allow for a genuine science of the artificial. We think a correct understanding of final causality will overcome current resistance to this principle.

Keywords: finality; mechanism; consciousness; theory of the four causes; science of the artificial.

1. Introduction: re-discovering finality

The birth of modern science in the Renaissance is marked by the rejection of all kinds of unverifiable explanations, especially explanations based on the principle of finality, which were often assimilated to mythical explanations (Cantor & Klein 1969). According to the principles of the experimental scientific method, we can discover the laws that govern the phenomena of nature, but we cannot discover any purpose or finality beyond those laws. Consequently, we can analyze the behavior of an artifact as a phenomenon, but we cannot discover its purpose based solely on observations: reverse engineering is beyond the capabilities of the experimental scientific method.

This perfectly legitimate methodological self-limitation of science, however, was adopted as a fundamental metaphysical principle by mechanistic philosophy, with its view that finality is a *subjective illusion*. There is no finality in nature, because —it is implicitly assumed— the only way to acquire objective knowledge about nature is through the experimental scientific method. In our view, this raises an interesting paradox in the realm of artifacts, since technology —which is intimately bound up with modern science— cannot be understood without the principle of finality. As we will see, this principle simply means that finality —i.e. having a purpose— is essential to understanding human artifacts.[1] Indeed artifacts are mainly defined by their purpose, which becomes their design principle and the guide to their construction. Therefore, if science excludes the principle of finality, then the realms of science and technology are divorced, because technology is not *objective knowledge*: technology cannot be scientifically studied because it does not belong to nature, it is subjective, i.e. projected from the mind onto material reality.

The relationship between science and technology has been studied for some time. In the last decades of the 20th century a growing consensus took shape: the scientific method, developed for studying and analyzing natural phenomena, was not suitable for understanding the design and construction of human artifacts [Génova et al. 2012]. The required method, it was argued, should start not with the observation of phenomena, but rather with the identification of a need, followed by artifact construction and evaluation [Hevner et al. 2004]. Probably the most prominent proponent of establishing research disciplines that do not follow the classical scientific model was Herbert Simon, who in his seminal work *The sciences of the artificial* [Simon, 1969] popularized the concepts of design sciences as opposed to natural sciences. However it is still often argued that design (technology) and science are not essentially different activities [Farrell & Hooker 2012]. This view has been convincingly refuted: it has been shown that design and science, even if closely related, are different kinds of intelligent activity so far as their aims, subject matter, products and methodology are concerned [Galle & Kroes 2014]. We would like to join this debate by discussing the role of the principle of finality in objective knowledge.

We think —and it is the claim we put forward in this paper— that philosophy of technology demands a paradigm shift in our notions of causality and objective/scientific knowledge. We are convinced that the association of the principle of finality with mythical explanations has led to a serious misunderstanding. Instead we think that a correct understanding of final causality can help us rediscover the value of the principle of finality in the explanation of the artifacts we produce with our technology. Such an understanding will be of particular relevance to the comprehension of computational machines and of artificial intelligence, and indirectly of the human mind. We think this rediscovery of finality is essential to a genuine *science of the artificial*, even if it takes us beyond the scope of the classical scientific method.

The rest of the paper is structured as follows. In sections 2 and 3 we show that the notion of machine —contrary to a widespread philosophical tradition— necessarily

[1] In this paper we are concerned only with finality in human artifacts and machines. In the context of human artifacts, finality can be assimilated to voluntary purpose, i.e. the purpose that has been intentionally implemented in the artifact by its artificer. We do not necessarily subscribe to the stronger form of the principle, according to which "every agent acts for the sake of an end" (*omne agens agit propter finem*, as Scholastic philosophy formulated it [Maritain 1934]), especially when it is implied that having finality implies being conscious, or having been consciously designed.

includes its having a purpose, but that this purpose cannot be discovered through the experimental scientific method alone. In sections 4 and 5 we argue that finality and structure are essential to understanding a mechanical device, yet they are not —obviously— physical components that interact with other physical components; finality and structure cannot be conceived as a strange kind of efficient cause, but they are necessary explanations anyway; analogously, consciousness is not a physical thing that interacts with the body and causes behavior, but it nevertheless explains behavior. In sections 6 and 7 we claim that the design and construction of artifacts —including, of course, the consideration of finality— is also scientific (i.e. objective) knowledge, and we try to clarify some misunderstandings regarding the Aristotelian theory of the four causes. We conclude in section 8 with final considerations about the misunderstanding of Aristotle's views on causality that has led in our tradition to the epistemological primacy of science-theory over creation-practice, overlooking the fact that it is the process of invention, design and production of human artifacts that provides Aristotle with his conceptual model of causality.

2. The notion of machine: a machine is not (only) a mechanical system

The concept of machine encompasses two different aspects: something subject to mechanical laws, and something designed for an end. However, the second aspect is very often forgotten, even denied, or it stirs argumentative reluctances which are not always explicit.[2] In general, the use of the term 'machine' to refer to a 'mechanical system' without reference to its end (or even denying that it has an end, especially when it is applied to living beings), is linked to Mechanism, a philosophical stance that has roots in Greek atomists like Leucippus and Democritus (5th cent. BC). Surprisingly, as we will see, many such mechanistic thinkers used the term 'machine' precisely in an attempt to avoid the necessity of a Creator. Mechanism developed in modern philosophy mainly through the works of René Descartes (1596-1650) and Thomas Hobbes (1588-1679), and it was famously formulated in conjunction with determinism by Pierre Simon de Laplace (1749-1827):

> *We may regard the present state of the universe as the effect of the past and the cause of the future. An intellect which at any given moment knew all of the forces that animate nature and the mutual positions of the beings that compose it, if this intellect were vast enough to submit the data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom; for such an intellect nothing could be uncertain and the future just like the past would be present before its eyes.* [Laplace 1814]

We have no wish to enter an academic dispute over definitions, especially given the cultural inertia that promotes the view that the natural sciences deal exclusively with physical actions. Simply we wish to argue that, in this notion of 'machine' as 'non-finalized mechanical system', an absolutely fundamental element is missing. According to this notion, which only considers purely physical or mechanical aspects, the Solar System is a machine. But is it? To us it is quite obvious that it is not. The Solar System

[2] See for example [Arana 2015] and [López Corredoira 2005] for attempts at naturalistic accounts of machines. See also [Génova 2016] for a previous version of our counter-argument.

is not a machine. Why? Because it has not been *designed*. Only if (a hypothesis that we may not want to rule out altogether) we think that the Solar System has been constructed by a powerful extraterrestrial civilization, which has calculated and established the size of the Sun and the orbits of the planets so that the habitable zone is a determined one and life on planet Earth can arise, and so on, then yes, the Solar System would be a machine, because it would have been designed with an end in view.

Therefore, if it has not been designed for an end, then the Solar System is not properly a machine, although metaphorically we may refer to it as such, perhaps as an inheritance of a tradition that still sees the cosmos and its elements as artifacts of its Creator (thus endowed with a "natural" purpose or finality). Gottfried Wilhelm von Leibniz (1646-1716) used the expression 'natural machines' as opposed to 'artificial machines' in his *New System of Nature* [Leibniz 1695], in an attempt to revive the Aristotelian distinction between animate and inanimate things, and resist the Cartesian assimilation of natural machines to artificial ones [Raymont 1998, Nachtomy 2011]. Whether living beings have or have not a natural finality is not the subject of this paper. In any case, the use of the expression 'natural machines' to refer to non-living mechanical systems — such as the Solar System— is even more inadequate. We think that to speak of 'natural machines' in the sense of 'natural mechanical systems' (inanimate as the Solar System, or animate as living beings), is now more inappropriate than in the past, when the idea that the Cosmos was the work of a Creator still held sway. Even referring to them as 'systems' is in a certain sense improper, because the idea of ordering, of purpose, is implicit in any 'system'.[3] Indeed, it is very difficult to think and speak about the real world and at the same time truly and radically exclude all purpose.

Thus if purpose and design are essential to the notion of machine, then using the term 'machine' to imply a 'non-finalized mechanical system' is equivocal. Furthermore, to say that "a machine is any conjunction of material elements which performs *exclusively physical actions*" [Arana 2015] is incomplete, and we dare say incorrect, since the omitted element is essential. Even more incorrect is the claim that the operation of the machine is explicable *only in terms of the mechanical interaction* of its components [López Corredoira 2005]. It seems to us that dictionaries are wiser when they say that a machine is "an aggregate of ordered parts and *directed* to the formation of a whole" [RAE 2017], "a material structure *designed* for a specific purpose" [OED 2017], "an instrument (as a lever) *designed* to transmit or modify the application of power, force, or motion" [MWD 2017], or "a tool containing one or more parts that uses energy to perform an *intended* action" [Wikipedia 2017; Tatnall & Davey 2016]. The key point here is indicated by the words 'directed', 'designed', and 'intended', each of which underlines the purposeful nature of the machine.[4]

[3] Dictionary definitions (see [OED 2017] and [MWD 2017]) of the English word 'system' refer primarily to artifacts or living beings in which function (ordering or purpose) is clearly implied; secondarily, it is also used for groups of objects in the context of material sciences such as Geology and Astronomy, where there is no purpose, and ordering is accidental. We can assume that the same happens in any modern language that has borrowed the term from Greek, where it literally means 'composition'. This is certainly the case in Spanish.

[4] Properly speaking, 'design' is a broader term than 'purpose'. Design encompasses purpose, but also structure, the choice of materials employed, etc.

Having a purpose does not necessarily imply being conscious. Julien-Offray de La Mettrie (1709-1751), a French physician and philosopher who claimed that human beings are mere automatons or machines, wondered in his famous *L'Homme Machine* [La Mettrie 1747] whether matter can think; whether, for example, a clock has any intentionality in telling the hours. Our answer would be that telling the hours is an *interpretation* of the movement of the clock, and that being an interpretation, it corresponds to a mental act that lies outside of the clock itself: the clock hands do not tell the hours, it is the one who looks at the clock, or the one who has made it, who does so. In the same way, a thermostat does not realize that it measures the temperature, but its manufacturer and its user do: this indeed is where we find the purpose, the intentionality, the meaning. In other words the actions of a machine are not *exclusively physical* nor can they be explained only as mechanical interactions: they are *intentional actions*, directed towards a purpose, which someone has established from outside the machine itself. Therefore it is impossible to understand a machine without understanding its purpose, but —at the same time— a machine need not be conscious to execute intentional actions that are meaningful to an external consciousness. We intend to illustrate this thesis with a simple and inexpensive mental experiment.

3. A mental experiment in the attic: purpose is not discovered with the scientific method alone

One of the typical tasks of engineering professionals is what we call *reverse engineering*. That is, given an artifact, you try to find out how it works and what purpose it serves, in order to improve it or, if necessary, imitate it. At first it may seem that the traditional experimental scientific method is adequate to performing this task, but a closer look reveals that this is not the case. In fact discovering *what an artifact is for*, in other words its purpose, is tantamount to discovering the *intention* for which it was made (or even the intention for which it is used, which may not coincide with the first: think of all the objects we use as paperweights...).

Consider the example of an artifact of uncertain appearance that we find in an old attic. We examine it closely, and find that it has a cyclical movement with a periodicity of 24 hours plus one minute, in other words, *it looks like a clock*. So far so good: this observation of a regular movement falls within the scope of the experimental scientific method. Now, from this observation alone we cannot come to a conclusion about *the intention* of the manufacturer who built such an artifact. What we can do is propose hypotheses about this intention, among which the following three might suggest themselves: (a) the purpose of the artifact was to display a cyclical movement covering a period of 24 hours, but it turns out that for whatever reason the period is somewhat greater, i.e. *there is a malfunction*; (b) the purpose of the artifact was to display a cyclical movement with a periodicity of 24 hours plus one minute, so it turns out that *it works perfectly*, for that is the periodicity displayed by the artifact; (c) the purpose of the artifact was one that had nothing to do with cyclical movement and which has not yet become apparent to us — the cyclical movement occurs purely by chance, or through some unintended effect of the manufacturing process.

The important point in terms of this essay is that *there is no imaginable experiment* that could serve to determine which one of these options is the right one. The only way to know for certain what the manufacturers intended, and therefore what purpose the device has, is *to ask the manufacturers themselves.* Or one might find out what they had in mind from the instructions manual, or from some other public expression of their intention. It might even happen that the artifact was conceived as a present or as a decorative motif, so its accuracy as a measurement of time is irrelevant.[5] From this we conclude that the experimental scientific method, in the strict sense, is *insufficient for finding out what an artifact is for*, that is, for performing reverse engineering, even if reverse engineering is a regular part of an engineer's work.

Hence we can conclude that it is necessary to abandon a paradigm of thought that *impoverishes* both science and engineering [Génova et al. 2012]. The experimental scientific method does not have the answer to *all* questions, and least of all to questions regarding the purpose or finality of the artifacts that engineers build. In the realm of artifacts, the human world of intentions is as real as the world of physical-mechanical relationships. To do good engineering, *it is not enough to master the laws of nature*; the hermeneutics of intentionality and the context of the symbolism of artifacts are also essential. In fact even to counterpose "a physical action" to "an intentional action" is misleading, for every human artifact implements both in the perfect compatibility of a *dual nature* [Kroes 2010]. The whole operation of a mechanical artifact consists in an arrangement of physical actions entirely governed by the laws of nature, and at the same time a system of intentional actions —i.e. aimed at a purpose— designed by the manufacturer: they are not two different sets of actions, but two different views on the same set of actions. Therefore, *it is impossible to have an adequate understanding of the operation and existence of a machine without understanding its purpose*.

By this we do not mean to say that the experimental scientific method is mistaken in its deliberate disregard of finality, but only that it is insufficient for a full understanding of those realities characterized by purpose, such as free acts, and the products of those free acts, such as human artifacts. The scientific method is legitimate insofar as it goes, in its discounting and setting aside of finality; what is not legitimate is to turn this into an absolute, and to identify the realm of objective/scientific knowledge with the realm of natural, physical-mechanical relationships. We do not intend to disavow the scientific method, but only to dethrone it from a place that does not belong to it. Or rather we claim, in the same sense suggested by Herbert Simon in his call for a true *science of the artificial* [Simon 1969], that designing and producing an artificial thing that really works in the world, but is not spontaneously provided by nature itself, is also an

[5] See Ruth Millikan's analysis of the concept of 'proper function', which she contrasts with other theories that attempt to elicit a function from the actual working or disposition of the component parts. Present properties or activities, which can be observed through the experimental scientific method, are not enough to determine function, precisely because they do not account for defective organs or artifacts. "Indeed, a thing that bears no resemblance to any can opener previously on earth —suppose it has been designed in accordance with a totally new principle— may still be a can opener, and may be one despite the fact that it doesn't work. (…) My claim has been that accounts of purpose or function in terms of present disposition or structure run afoul exactly when they confront the most central issue of all, namely, the problem of what failure of purpose and defectiveness are" [Millikan 1989]. The concept of 'proper function' is broader than voluntary purpose (see note 1), since it encompasses also the finality of bodily organs.

objective (i.e. scientific) form of knowledge, not only about this thing but also about the world itself. It is *true science*, even if it departs from the classical experimental model of science in that it includes the consideration of finality with regard to artifacts.

The fact that artifacts have a purpose is something obvious and admitted without difficulty even by people with as little philosophical baggage as engineers tend to have. The real point, however, is that purpose is not discoverable through the experimental scientific method, understood in the strict sense, because *finalistic explanations are neither verifiable nor falsifiable*. Scientific experiments can tell us what happens, and how regularly. They can rigorously establish regularities and laws. But they cannot assure us that those regularities respond to an intentional design, nor what that design was. Indeed there is no way of knowing with certainty about intention other than by questioning the author. Properly speaking, there can be no *empirical* evidence that design exists.

This is even clearer in computational machines, i.e. machines that follow mechanical rules of behavior written in a program. An algorithmic computation is not simply something that *happens*, but something that happens *with a purpose*, i.e. it does something *for somebody*. Having a purpose is an essential element in the definition of algorithm [Hill 2016]. If comments in programming are so important it is because they declare the programmer's intention with regard to a fragment of code, something that is not so easy to grasp from the inspection of the program code itself, or from the observation of the program execution. In fact, this has been theoretically confirmed: we know that we cannot algorithmically determine whether a sequence of observed events (such as the behavior of an electronic device) has or has not a purpose, as Gregory Chaitin demonstrated in a derivation from Turing's Halting Problem [Chaitin 2005].

4. Finality and structure of machines: to see a thing you must first understand it

One of the most important starting points of Renaissance thought is its rejection of the principle of finality in the understanding of the universe: the search for a finalistic explanation is sterile, because it does not really explain what it claims to explain. In the oft-quoted words of Francis Bacon: *nam causarum finalium inquisitio sterilis est, et, tanquam virgo Deo consecrata, nihil parit*.[6] Perhaps it is true that the principle of finality is sterile when it comes to understanding *what actually happens* in nature; however, that same principle is essential to understanding *what is required to happen* in human artifacts. A machine that simply "does things" (physical actions) is not properly a machine. If the machine does not do what I want, the machine does not work well. It is its success or failure in performing its function that permits us to tell whether a machine works properly or not (quality control). Therefore, *a machine cannot be defined and accounted for without reference to its purpose*.

It is often said that the experimental scientific method and the naturalist perspective leave out art, metaphysics, ethics and religion. Well, we think we have shown that they

[6] *De Augmentis Scientiarum*, III, 5: "the inquiry of final causes is a barren thing, or as a virgin consecrated to God [it gives birth to nothing]" [Bacon 1623].

also leave out engineering and technology, since, we insist, it is impossible to understand a machine without understanding its purpose, its finality, its function. And, as it happens, it is also impossible to understand a machine without understanding its structure.[7] If engineers do anything, it is to design the purpose and structure of machines [Kroes 2010]. Likewise if there is one thing that computer science students are expected to learn, it is surely the design of computational machines and information systems with a view to their finality and structure.

There is an obvious parallel here with Aristotle's final and formal causes, which incidentally he explains in a technological context: the carpenter who makes a wooden table, or the craftsman who makes a statue.[8] We are not bent on reviving Aristotle, and we do not care too much if his causes were three, four or five, or if they are called one way or another. We think that fighting over words is not too sensible. Nonetheless it is curious, even paradoxical, that a philosophy should be called "mechanistic" which, from its first appearance until the present day, has dispensed with those very final causes without which mechanisms cannot be understood. To repeat, a mechanism, indeed any artifact, cannot be understood without understanding what it is for. As one great writer says, whose words the reader may recall:

> *In order truly to see a thing, one must first understand it. An armchair implies the human body, its joints and members; scissors, the act of cutting. What can be told from a lamp, or an automobile? The savage cannot really perceive the missionary's Bible; the passenger does not see the same ship's rigging as the crew* [Borges, 1975].

To see a thing you must first understand it. And to understand it is to understand what it is for, its purpose or finality. Focusing on computer theory, this is a perfectly accepted principle, which was already established by Alan Turing and the founders of computer science even before the advent of electronics [Turing 1936]: an algorithm (or *effective computation*) is a rule-based procedure that obtains *a desired result* in a finite number of steps. That is, one of the essential elements of the algorithm definition is its purpose, the desired goal or outcome that it has to pursue [Hill 2016]. By contrast, the material aspect is secondary, so that a computational machine can be implemented even by humans performing calculations: "A man provided with paper, pencil, and rubber, and subject to strict discipline, is in effect a universal machine" [Turing 1948].

In our view, the interesting thing about this approach is that finality is not a physical component of the machine: finality does not *interact* with the other components. The finality of the clock does not interact with the hands. The finality does not exert any kind of force by contact, or at a distance, with the elements of the machine; and, of course, the structure does not interact with them either. And yet, as we have already stated, finality and structure are essential to understanding what a machine is in general, and what a given machine is in particular. Finality (and structure) *influence* the operation of an artifact, but obviously do not physically *interact* with its elements.

---

[7] Here 'structure' means simply a coherent disposition or arrangement of the parts of a whole (what a software engineer would call 'architecture'); 'finality' means simply 'purpose', see note 1.

[8] See *Physics* II, 3 and *Metaphysics* V, 2 [Aristotle d. 323 BC].

Something analogous happens in a very particular way in computational machines, with their well-known stratification into hardware and software, the physical level and the logical level. Software is not an efficient cause of hardware behavior. Software is not physical, so it cannot *interact* with hardware. And yet no computer scientist will have the slightest doubt that software *influences* in a very real way the hardware. How are we to think of that influence? Using very old language, we can say that software is not an efficient cause but a formal cause: software is itself in*form*ation, interpretation. Software is *the-name-we-give-to* what happens in hardware: it is an *interpretation* of physical-mechanical interactions.

We think this throws an interesting light on the problem of the relationship between mind and body: of course *the mind does not interact with the body, in a physical sense*. And yet the mind is indispensable for understanding the thinking being as a whole. The mind is not an additional *component* of the human being, but its *intentional dimension or aspect*, without which it is impossible to understand it fully, as is the case with the finality and structure of a machine. We think this view is closely connected to the double-aspect theory that has been revived in recent decades [Chalmers 1996].

5. Cause and explanation: consciousness is not the efficient cause of behavior

In Aristotelian language we would say that the finality of an artifact is the *cause* of its operation. While this formulation may be acceptable, it currently runs the risk of being misunderstood. Since finality is not physical, nor does it interact physically, it cannot be an efficient cause. This would be obvious to anyone familiar with Aristotelian language: finality refers not to efficient cause, but, precisely, to final cause. But for those who do not master this language, the expression may become misleading. If finality is not an efficient cause, it might be argued, then it is no kind of cause, given the modern elimination of finality as an acceptable causal category.

That is why it seems more pertinent to speak of *explanation* instead of *cause*, while respecting the original meaning of the Greek term used by Aristotle (αἰτία, *aitia*), which today is obscured when it is translated univocally by the Latin term 'cause' [Hankinson 1998, Natali 2013, Falcon 2015], precisely because of the modern semantic reduction of the latter to 'efficient cause'. Therefore, rather than cause, it is far better in modern languages to say that finality is the *explanation* of the operation of the artifact. This makes it much easier to understand, and conveys better what is meant.

For the same reason, we propose to avoid saying that consciousness, or the mind, is the cause of behavior, as an observable phenomenon, but rather that consciousness explains behavior. *Consciousness is not (efficient) cause, but explanation.* In other words, consciousness is not a cause in the Cartesian sense of efficient causality, but it is cause (explanation) in the Aristotelian sense of final or formal cause. To insist on saying that consciousness causes behavior, very much in the Cartesian dualistic line, slides dangerously onto a terrain we do not wish to tread, for it leads directly to the attempt to convert consciousness into another observable, natural phenomenon; the attempt to convert it into a "thing" that supposedly "interacts" with something else, that other thing we call the body. In short, it is the attempt to naturalize consciousness, to give it a fully naturalistic explanation [Dennet 1991]. All this confusion can be avoided simply by saying that consciousness is not a cause, but an explanation.

This provides us with a suggestive clue as to *why human consciousness remains unexplained*: for *consciousness is not understood as explained* by something else (still less, explained by efficient causes), *but is seen instead as an intentional explanation* of specifically human behavior. To understand consciousness one must try not to explain it, but to understand what it explains (i.e. behavior), and thus how consciousness explains behavior.

So, what is the difference between machines and humans, between software and mind? The characteristic of a machine is precisely that it has purpose, a purpose conferred on it *from outside* itself. By contrast, we human beings are characterized by our ability to propose *our own* ends. We can come up with plans, projects, vital objectives; *we can decide what we want to be in life*. Thus human beings are, at least to some extent, self-determined. But, note, not self-designed: we are ourselves, in a very appropriate sense, natural beings, not artifacts. An artificial machine, on the other hand, and in particular an algorithmic or computational machine, cannot decide what goals it wants to pursue, because then it would by definition cease to be a machine. We consider that this ability to self-propose our own ends is the most characteristic difference between humans and machines. Perhaps in an indeterminate future we will be capable of producing in the laboratory a kind of non-algorithmic robot (not directed towards a given end) that properly can be described as self-conscious, able to do "whatever it wants", to propose its own objectives; but surely it would not be appropriate to continue calling such a thing a robot. It would be an entity with *artificial consciousness*, but it would not be a *computational machine*, since it would not be governed by a given and extrinsic purpose or finality. Self-determination, by its very definition, falls outside the classical computational paradigm and the general concept of machine.[9]

6. Philosophy of technology: a necessary paradigm shift in our notions of causality and objective/scientific knowledge

In the history of Western philosophy the three most influential perspectives on the nature of human artifacts have been, perhaps, those of Aristotle, Marx, and Heidegger [Mitcham 1994]. All of them base the power of their argument on an analysis of causality in general, and of the kind of causality that should be attributed to artifacts and to human production. It is out of this implicit *context of production* that the *theory of the four causes* in Aristotle arises,[10] together with the development of *historical materialism* as a genuine causality of history in Karl Marx, and the notion of *enframing* or *gathering-together (Gestell)* in Martin Heidegger [Heidegger 1954, Godzinski 2005]. Now in both Heidegger and Marx the notion of cause as such, and its function as an explanatory argument, are based in their fundamental lines on the scheme proposed by

[9] We have extensively developed this idea in a different work [Génova & Quintanilla Navarro 2018].

[10] A general account of the doctrine of the four causes is found, in almost the same words, in *Physics* II, 3 and *Metaphysics* V, 2 [Aristotle d. 323 BC, Falcon 2015].

Aristotle and on the idea that knowledge by causes defines a special type of knowledge that we call scientific.

But whether we call them 'causes' or 'explanations', there is an underlying need for a true paradigm shift in our notion of scientific (i.e. objective) knowledge insofar as it pertains to the causality of technology in general, and the causality of machines in particular. The mechanistic model of the universe, of mind and of human action that has dominated the modern period of philosophy has been already *de facto* disproved, precisely, by technology. As was the case with the original Industrial Revolution, post-industrial technological development has very significant theoretical repercussions on this philosophical debate. These technological developments directly affect our theoretical discourse on causality, on which our notion of scientific knowledge still depends. In particular, the attribution of final causality to human artifacts raises wide-ranging theoretical implications. Herbert Simon anticipated the scope of these implications by advocating a genuine *science of the artificial* [Simon 1969].

However the weight of our tradition of thought continues to restrict the scope of true causality to that of an *efficient* and *natural* causality; if artifacts have a purpose, it is only because *we attribute it* to them through the way we use them; there is no essential difference between building a walking stick and using a casually found branch to walk. In a more extreme example, there would be no essential difference between using animal horse power and steam machine 'horse' power. Certainly the laws of thermodynamics and their particular form of causality are found in nature, but the power of a steam machine depends not only on them, but also on the structure and purpose with which the machine has been built. Equally, electromagnetism is natural, but the operation of an algorithmic computer cannot be explained, i.e. *caused*, by electromagnetism alone. This metaphysical restriction of causality to efficient causality contradicts the immediate experience that artifacts possess true final causality (i.e. design, purpose), beyond the actual use of them, and it restricts finality to the realm of *subjective* attribution, making a genuine science of the artificial impossible, because artificial things are —supposedly— not *objective*.

In the current systemic model of technology, the prevalence of the unit of information over the unit of mass in our rationalization of the universe obliges us to rethink the entire Western philosophical discourse about material, efficient, formal and final causality. Immanuel Kant's clear-sighted intelligence was fully aware of this problem and tried to prevent the ascription of objective character to formal and final causes by closing the road to a genuine attribution of causality to technology: that is, by denying that artifacts actually had an *objective purpose*. In Kant's view, human technique is not a true dimension of change in the universe, it is not an objective process, and the artifact brings nothing to the causality of nature; thus he presents a conception of nature (encompassing physical artifacts considered as mere physical objects) in which we cannot properly speak of finality, in conformity with his entire philosophical system. Hence Kant promotes the notion of 'usability' as a key category for understanding the causal action of the human artifact in the form of *subjective attribution of purpose* [Quintanilla Navarro 1998], as we have explained with the examples of the walking stick, the steam machine and the algorithmic computer.

Anticipating Herbert Simon by several years, the German engineer and philosopher Friedrich Dessauer challenged this Kantian unnecessary counterposition between the artifact as a physical and a teleological (i.e. finalized) object, and he tried to develop an alternative conceptual model [Dessauer 1956]. He pointed out, among other things, that a mechanistic metaphysics preventing the development of a genuine science of the artificial can arise only on the basis of an *a priori* definition of nature (φύσις, *physis*) as the exclusive source of causality —i.e. not only as a necessary instance, but as the only possible instance— and in addition when we *a priori* restrict the action of nature to efficient and material causality.

We could say that the epistemological model inherited from the nineteenth century implicitly assumes not only that no natural object or process possesses any purpose of its own, but that there can be no material system in which this type of causality operates objectively. *This implicit assumption makes human freedom something tremendously paradoxical and problematic for philosophy*, as the Nineteenth Century soon came to understand: if we humans are mere natural beings, i.e. the products of a natural process without finality... how is it possible that we are capable of building finalized artifacts and, more important, how is it possible that we human beings self-determine ourselves towards the ends we choose? All the fundamental effort of Kantian thought is oriented towards recomposing the rationality of human action by assuming the *insurmountable abyss* between the physical and the intentional worlds that the new notion of nature entailed.

But freedom and consciousness are not the only aspects of our radical experience of reality that are expelled from a physical universe which has become the purported exclusive domain of objective knowledge. Together with freedom and consciousness, technology itself is also banished from the realm of objective knowledge, since technological causality —as the producer of purposeful physical objects and processes— is, in Kant's view, a subjective illusion. If there is an *objective* causality, so the argument goes, then it must be *natural* causality and, therefore, *unintentional*. Within these epistemological parameters, artificial intelligence is an impossible chimera: nothing material could be intelligent because there is an effective equivalence between all that is material, objective, natural and unintentional, and therefore nothing unintentional could be really (objectively) intelligent. Consequently, the implementation of intelligence in a computer is not objectively real; the only objective reality found within it is the series of physical processes (mechanical, electronic, or whatever); interpreting those phenomena as an intelligent or intentional computation is only a subjective attribution of finality.[11] To this we reply that, certainly, finality is not *verifiable* using the classical scientific method alone, as we have illustrated with our mental experiment in section 3. But, in our view, this does not mean that the finality of artifacts is not *real*; rather it demands a broadening of the modern notion of science to encompass the design and construction of artifacts, in other words a *science of the artificial*.[12]

[11] Note that this implies, in Kant's conception of humanity, that we are ourselves incomprehensible as simultaneously material and rational beings. *De nobis ipsis silemus* (let's keep silent about ourselves) are the words of Bacon with which Kant starts his *Critique of Pure Reason*.

[12] Note that this broadening of science to encompass artificial finalities does not imply that 'natural finalities' can be scientifically studied, too.

7. The contemporary value of the theory of the four causes: four misunderstandings

In one of the few references to technology found in Karl Popper's work, it is pointed out that *each artifact by itself refutes a theory*: the one in which the artifact is impossible [Popper & Kreuzer 1986]. This implies that the universe is an open system, a structure capable of taking on goals that are not determined from the start: either self-determined goals (in humans) or at least hetero-determined goals (in artifacts). Therefore, the very existence of artifacts refutes the theory that there are no purposes in the universe —the mechanistic model of the universe— because each artifact is obviously a finalized structure. It could be argued that those ends are somewhat "above" the material universe, yet this only goes to show the need for considering and distinguishing the different levels of causality (i.e. explanation). Besides, "being above" does not mean "not being real", unless one *a priori* assumes that the only real aspect of the universe is matter, and that the only way of being real is being material.

Human technology in general, and artificial intelligence in particular, show that what we call the material universe is a structure capable of being shaped towards ends — at least *our* ends. And yet this evidence meets with enormous resistance within much of our philosophical and epistemological tradition, which continues to accept the postulate that teleological action is always and exclusively subjective —that is, attributed by a conscious agent— and cannot be algorithmically expressed, in the sense of being objectively "implemented" as a mechanism in a material body. In other words if goals are free, it is argued, then they cannot have a material substrate, because matter is not capable of "receiving" goals: matter and freedom are divorced. This postulate renders inconceivable —i.e. impossible to understand, impossible to be— not only free human beings, but also freely designed artifacts, as Dessauer had pointed out.

The logical analysis of a number of explicit or implicit arguments on this question reveals that, when we speak of the true causality that opens up the understanding of an artifact, be it an arrow, a piano or a computer, we often succumb to misunderstandings, which should therefore be made explicit:

1. The **first misunderstanding** is essentially historical and consists in thinking that, in the Aristotelian model of the four causes, it is theoretically possible for only one of them to produce and explain the system or process independently of the others. This is a mistake. What makes something come into being, and what makes something act as it does, and what makes us really know what something is, are three different issues that human reason tries to understand through causal explanations. Often, however, in our causal explanations we jump from one level to another without precise argumentation. In Aristotle's theory of causality, these three different issues are closely related, they even have an essential and inescapable linkage, but they are not the same. Yet in all three levels of argument, it is necessary to appeal to a material cause, an efficient cause, a formal cause and a final cause. The four concur in a single action and the four shape a single explanation. The modern Baconian mutation that scientifically

explaining is being able to produce,[13] does not really justify the conclusion that material and efficient causality provides a sufficient explanation.

2. The **second misunderstanding** is to assume that the final cause in the activity of an agent is the conscious intention in such an agent. Freedom, consciousness, intelligence and intentionality are closely interconnected, but they are not synonymous. Having a purpose does not necessarily imply being conscious, as every human-made artifact demonstrates. There are two forms of determination with regard to an end: self-determination and hetero-determination. Only the former requires freedom and consciousness in the agent (a human); the latter can be satisfied with an extrinsic finality, which does not spontaneously originate in the agent (an artifact), but which is nevertheless really "implemented" in it, so that this hetero-determination truly belongs to the artificial agent and conforms its "essence". The finality of a programmed robot is not in its intentions (which, not being conscious, it does not have), but in the intentions of its programmer; a robot, like any machine, has a purpose, but not an intention. Both intention and purpose are legitimate forms of final causality, either conscious or unconscious. Besides, note that the programmer who implements the purpose is the efficient cause of the robot, not its final cause; the final cause is the implemented purpose, not the programmer who implemented it.

3. The **third misunderstanding** refers to the very relations between knowledge and production. As we have seen, when we talk about the purpose of a clock or a computer program we are tempted to think that *reverse engineering* is the only way to obtain and justify objective/scientific knowledge about the world. Now, *forward engineering* (to produce something or to make something effectively work) is also knowledge about the world. It is a genuine knowledge about what works and does not work, what can and cannot be done. Denying the rational "certainty" of the objective purpose of our inventions is a poor way of approaching objective knowledge as it expresses itself in the relationship between the human being and the universe.

4. The **fourth misunderstanding**, finally, is to imagine that by sticking to a mechanistic model based on efficient and material causality we do not take on any metaphysical presuppositions, whilst the reference to finality in natural beings compels us to philosophize about a putative creator of the universe. But this implicit claim on "metaphysical neutrality" does not stand up to examination. On the contrary, one of the main disciples of Descartes, Nicolas Malebranche (1638-1715), who was a fervent Christian believer, argued quite convincingly that it is precisely a mechanistic conception of the universe that requires an Artificer to rationally answer the question of its existence, whereas a universe such as the Aristotelian —pervaded with intrinsic causations— is the one that can be explained in completely immanent terms, without reference to its

[13] In the words of Francis Bacon, *Novum Organum*, I, 3: "Knowledge and human power are synonymous, since the ignorance of the cause frustrates the effect" (*Scientia et potentia humana in idem coincidunt, quia ignoratio causae destituit effectum*) [Bacon 1620]. That is, knowing (explaining) the cause is the same as being able to produce the effect.

> Creator. Malebranche considered this the most dangerous error of the ancient pagan philosophers [Malebranche 1674: VI, II, III].

In other words, there is no less metaphysics in the Cartesian conception of the human body as an automaton [Descartes 1648] than in Goethe's *Urpflantze* (his archetypical plant) [Goethe 1790]. The relation between the two notions of causality in the physical world —Descartes' mechanistic and Goethe's organicist outlook — and other notions such as freedom, the human soul or God, always and inexorably demand the same complex and controversial philosophical arguments. It is not possible to move from empirical science to the sense of the whole without going through metaphysical speculation.

8. Conclusion: understanding causality in a context of technological production

The theoretical analysis of technology as a kind of rational knowledge in Aristotle stops short of a true science of production, i.e. of the process by which *those things that could exist or not depending on the will of the artificer* come into being. In spite of this, there is no doubt that the conceptual framework of the four causes —material, formal, efficient and final— has been the most influential and consensual when analyzing the nature of our technical products. This conceptual framework —with necessary rectifications due to the development of science and technology— is not only still valid but also benefits from the new impetus that systemics and information technologies provide.

For example, the assimilation of *hardware to material cause* —always relative to the corresponding level of analysis because every material element of a system is, in turn, a complete system at another level of inferior analysis— as well as of *software to formal cause*, is not only a lucky heuristic coincidence: it is rigorously accurate in a proper and explicit sense. Of course, we do not claim that Aristotle anticipated the development of new computational technologies — notwithstanding the famous glimpse of the automatic servant.[14] Our point is that, right from the beginning, *it was the process of invention, design and production of human artifacts that provided in the Greek culture of Aristotle the conceptual model of causality.* And it was through that model that he defines "being", scientific knowledge, and the nature of that "first science" which will later be known as metaphysics.

As summarized by one of his best medieval interpreters —Averroes, Ibn Rushd (1126-1198)— "the activity of philosophy is nothing more than reflection upon existing things and consideration of them insofar as they are an indication of the Artisan — I mean insofar as they are artifacts" [Ibn Rushd d. 1198]. The theory of the four causes is already covert meta-engineering, and the concepts of nature, matter or efficient causality really derive from an archaic state of reflection on the basis of the artifact and

[14] Aristotle, who justified the need for slavery in his time, speculated in his *Politics*, I, 4, that automatons might someday make possible the abolition of slavery: "for if every tool could perform its own work when ordered, or by seeing what to do in advance (…), if thus shuttles wove and quills played harps of themselves, master-craftsmen would have no need of assistants and masters no need of slaves" [Aristotle d. 323 BC].

about the artifact. However, it seems that neither Aristotle nor the later tradition realized the centrality of the productive framework underlying his theory of the four causes; instead, the conception that science is "theoretical knowledge" triumphed, and consolidated the epistemological primacy of science-theory over creation-practice, an epistemological primacy which represents, in fact, an enormous misunderstanding of causality as propounded by Aristotle. The rectification of this misunderstanding motivates, among others, the reflections of Marx or Heidegger on technology, artifact and causality.

References

1. Arana, J. (2015). *La conciencia inexplicada*. Madrid, Biblioteca Nueva.
2. Aristotle. (d. 323 BC). *Aristotle in 23 Volumes*. Cambridge: Harvard University Press, 1944. (http://www.perseus.tufts.edu/hopper/).
3. Bacon, F. (1620). *Novum Organum*. Translated and edited by J. Devey: *Novum Organum; Or, True Suggestions for the Interpretation of Nature*. New York: P. F. Collier & Son, 1902.
4. Bacon, F. (1623). *De Augmentis Scientiarum*. Translated and edited by H. Morley: *The Advancement of Learning*. London: Cassell & Company, 1893.
5. Borges, J.L. (1975). *There Are More Things*. In *El libro de arena*. Buenos Aires: Emecé.
6. Cantor, N.F. & Klein, P.L. (1969). *Seventeenth-Century Rationalism: Bacon and Descartes*. Waltham, Massachusetts: Blaisdell.
7. Chaitin, G. (2005). *Meta Math! The Quest for Omega*. New York: Vintage Books.
8. Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
9. Dennett, D. (1991). *Consciousness Explained*. London: The Penguin Press.
10. Descartes, R. (1648). *Traité de l´Homme*. Edited by S. Gaukroger: *Descartes: The World and Other Writings*. Cambridge: Cambridge University Press, 1998.
11. Dessauer, F. (1956). *Streit um die Technik*. Frankfurt am Main: Knecht.
12. Falcon, A. (2015). *Aristotle on Causality*. The Stanford Encyclopedia of Philosophy (Spring 2015 Edition), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/spr2015/entries/aristotle-causality.
13. Farrell, R. & Hooker, C. (2012). *The Simon-Kroes model of technical artifacts and the distinction between science and design*. Design Studies 33(5):480–495.
14. Galle, P. & Kroes P. (2014). *Science and design: Identical twins?* Design Studies 35(3):201–231.
15. Génova, G., Llorens, J., Morato, J. (2012). *Software engineering research: the need to strengthen and broaden the classical scientific method*. In M. Mora, O. Gelman, A.L. Steenkamp and M. Raisinghani (eds.), *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems*. IGI Global, 2012, 106-125.
16. Génova, G. (2016). *Máquinas computacionales y conciencia artificial*. Naturaleza y Libertad 7:123–143.
17. Génova, G. & Quintanilla Navarro, I. (2018). *Are Human Beings Humean Robots?* Journal of Experimental & Theoretical Artificial Intelligence, 30(1):177–186.
18. Godzinski, R. (2005). *(En)Framing Heidegger's Philosophy of Technology*. Essays in Philosophy, 6(1):1-9, article 9.
19. Goethe, J.W. (1790). *Versuch die Metamorphose der Pflanzen zu erklären*. Translated and edited by D. Miller: *The Metamorphosis of Plants*. Cambridge: MIT Press, 2009.
20. Hankinson, J.R. (1998). *Cause and Explanation in Ancient Greek Thought*. Oxford: Oxford University Press.
21. Heidegger, M. (1954). *Die Frage nach der Technik*, in *Vorträge und Aufsätze*. English translation: *The Question Concerning Technology and Other Essays*. New York: Harper, 1977.
22. Hevner, A.R., March S.T., Park, J. & Ram, S. (2004). *Design science in information systems research*. MIS Quarterly, 28(1):75–105.
23. Hill, R.K. (2016). *What an algorithm is*. Philosophy & Technology 29(1):35-59.
24. Ibn Rushd. (d. 1198). *The Decisive Treatise Determining the Connection Between the Law and Wisdom, and Epistle Dedicatory*. Translated and edited by C.E. Butterworth. Provo: Brigham Young University Press, 2001.

25. Kroes, P. (2010). *Engineering and the dual nature of technical artefacts*. Cambridge Journal of Economics 34(1):51-62.
26. La Mettrie, J.O. (1747). *L'Homme Machine*. Translated and edited by A. Thomson: *Machine Man and Other Writings*. Cambridge: Cambridge University Press, 1996.
27. Laplace, P.S. (1814). *Essai philosophique sur les probabilités*. Translated from the French 6th ed. (1840) by F.W. Truscott & F.L. Emory: *A Philosophical Essay on Probabilities*. London: Chapman & Hall, 1902.
28. Leibniz, G.W. (1695). *Système nouveau de la nature et de la communication des substances, aussi bien que de l'union qu'il y a entre l'âme et le corps*. Translated from French by A.E. Kroeger: *New System of Nature, and of the Communication of Substances, as well as of the Union of Soul and Body*. The Journal of Speculative Philosophy 5(3):209-219, July 1871.
29. López Corredoira, M. (2005). *Somos fragmentos de Naturaleza arrastrados por sus leyes*. Madrid, Vision Net.
30. Malebranche, N. (1674). *De la recherche de la vérité et Éclaircissements*. Translated and edited by T.M. Lennon & P.J. Olscamp: *The Search after Truth and Elucidations*. Cambridge: Cambridge University Press, 1997.
31. Maritain, J. (1934). *Sept leçons sur l'être et les premiers principes de la raison spéculative*. Paris: Téqui. (*A Preface to Metaphysics: Seven Lectures on Being*. New York and London: Sheed and Ward, 1939.)
32. Millikan, R.G. (1989). *In defense of proper functions*. Philosophy of Science 56(2):288-302.
33. Mitcham, C. (1994). *Thinking Through Technology: The Path Between Engineering and Philosophy*. Chicago: University of Chicago Press.
34. MWD. (2017). Merrian-Webster Dictionary (https://www.merriam-webster.com/). *Machine*. *System*. Retrieved 01/31/2017.
35. Nachtomy, O. (2011). *Leibniz on Artificial and Natural Machines: Or What It Means to Remain a Machine to the Least of Its Parts*. In J.E.H. Smith, O. Nachtomy (eds.), *Machines of Nature and Corporeal Substances in Leibniz*, The New Synthese Historical Library 67, pp. 61-80, DOI 10.1007/978-94-007-0041-3_5, Springer Science+Business Media B.V. 2011.
36. Natali, C. (2013). *Aitia in Plato and Aristotle. From Everyday Language to Technical Vocabulary*. In C. Viano, C. Natali, and M. Zingano (eds.), *Aitia I: Les quatre causes d'Aristote: origines et interpretation*, Leuven: Peeters, pp. 39–73.
37. OED. (2017). Oxford English Dictionary (http://www.oed.com/). *Machine*. *System*. Retrieved 01/31/2017.
38. Popper, K.R. & Kreuzer, F. (1986). *Offene Gesellschaft - Offenes Universum, ein Gespräch über das Lebenswerk des Philosophen*. München: Piper.
39. Quintanilla Navarro, I. (1998). *Tecnología y metafísica: ¿hacia el final de una era kantiana?* Diálogo Filosófico 40:27–44.
40. RAE. (2017). Diccionario de la Real Academia Española (http://dle.rae.es/). *Máquina*. Retrieved 01/31/2017.
41. Raymont, P. (1998). *Leibniz's Distinction Between Natural and Artificial Machines*. Twentieth World Congress of Philosophy. Boston, MA, August, 1998. http://www.bu.edu/wcp/Papers/Mode/ModeRaym.htm
42. Simon, H.A. (1969). *The Sciences of the Artificial*. Cambridge, Massachusetts: MIT Press.
43. Tatnall, A. & Davey, B. (2016). *Towards Machine Independence: From Mechanically Programmed Devices to the Internet of Things*. In Arthur Tatnall, Christopher Leslie (eds.), International Communities of Invention and Innovation, IFIP WG 9.7 International Conference on the History of Computing, HC 2016, Brooklyn, NY, USA, May 25-29, 2016, Revised Selected Papers, vol. 491 of the series IFIP Advances in Information and Communication Technology, pp 87-100.
44. Turing, A.M. (1936). *On computable numbers, with an application to the Entscheidungsproblem.* Proceedings of the London Mathematical Society 2(42):230–265.
45. Turing, A.M. (1948). *Intelligent Machinery*. National Physical Laboratory Report, 1948. In Meltzer, B., Michie, D. (eds), *Machine Intelligence* 5, Edinburgh University Press, 1969.
46. Wikipedia. (2017). Wikipedia (https://en.wikipedia.org/). *Machine*. Retrieved 01/31/2017.

Biographies

Gonzalo GÉNOVA is Associate Professor of Software Engineering at the Computer Science and Engineering Department of the Universidad Carlos III de Madrid. He has an MS degree in Telecommunications Engineering, an MS degree in Philosophy and a PhD in Computer Science and Engineering. His main research subjects are models and modeling languages in software engineering, requirements engineering and philosophy of information systems.

Ignacio QUINTANILLA NAVARRO has a PhD in Philosophy and an MS degree in Psychology. He is currently Dean of the Infanta Elena High School in Galapagar near Madrid, where he teaches as a secondary school teacher, as well as part time lecturer in the Department of Theory and History of Education at Universidad Complutense de Madrid. His areas of research are the history of modern thought and the cultural and educational impact of our new technological environment, topics on which he has published various articles and monographs.