Will Machinic Art Lay Beyond Our Ability to Understand It?

Rui Penha, Miguel Carvalhais

INESC TEC / Faculty of Engineering, University of Porto; INESC TEC / Faculty of Fine Arts, University of Porto
Porto, Portugal
ruipenha@fe.up.pt, mcarvalhais@fba.up.pt

Abstract

In this paper we will argue that artistic creations made by artificial minds will most likely lay beyond our ability to understand them. We will assume that the emergence of consciousness in artificial minds is possible and that the artistic creations we are referring to are made by the artificial minds' own volition. We will build upon the definition of art as embodied meaning and explore its relationship with embodied cognition to argue that there is a binding of human artistic creation to the subjective experience of existing in a natural and cultural world through a human body that is born with a foretold death. Additionally, we will try to show that the best we can aim at, as human beings standing by an artistic creation by another species, is to an understanding of what could have motivated another human being to create such a work. As such, we shouldn't be able to understand an artistic creation originating by an artificial mind with a physical experience of the world that differs from our own, even if they have a privileged access to our culture. The boundaries for this incomprehension are those of the human mind.

Keywords

Machinic Art, Artificial Creativity, Artificial Intelligence, Consciousness, Aesthetics, Artificial Aesthetics, Definition of Art, Embodied Meaning, Embodied Cognition, Anthropomorphism.

Introduction: Four Assumptions

Our aim with this paper is to discuss artistic creation by artificial minds. We will question how it is that we, human beings endowed with natural minds, may or may not be able to understand, enjoy, empathise with, or even recognise machinic creations as art.

In order to develop this argument, we will make four assumptions:

First assumption: Sometime in the future artificial minds with at least human-level artificial intelligence will be developed (Tegmark 2017).

Second assumption: Consciousness is an emergent property (Gazzaniga 2011), it is substrate-independent and not contingent on any properties of human wetware (Hofstadter 2007).

Third assumption: The emergence of consciousness in artificial minds is possible, maybe even inevitable. We are very aware of the intense, long-standing, and continuing debate surrounding these questions, both in the fields of artificial intelligence and cognitive sciences (e.g., Turing 1950; Turing 1951; Bateson 1979; Pinker 1999; Eagleman

2011; DiSalvo 2011; Humphrey 2011; Oliveira 2017; Damásio 2017), as well as in philosophy (e.g., Jefferson 1949; Anderson 1964; Searle 1981; Penrose 1989; De Landa 1991; Dennett 1991; Bostrom, 2014) and in artistic and creativity studies (e.g., Boden 2004; Cope 2005; Hayles 2005; Ariza 2009; Deutsch 2011). In fact, even amongst the authors of this paper, the discussion is far from settled, but this fact is, as are these discussions, somewhat irrelevant to the arguments being proposed.[1]

We will argue that consciousness-endowed artificial minds will significantly differ from human minds or other organic and biologically evolved minds. And that therefore, their conscious experiences—because they are embodied (Metzinger 2009, 18-19)—will likewise be significantly different.

> The most important thing to know about thinking machines is that they will think different. (Kelly 2016)

Fourth assumption: Because we are discussing artificial intelligences that are at least human-level, that are endowed with consciousness, and therefore with comparable autonomy and agency, we will assume that these minds may also feel the need to engage in artistic activities.

When discussing machinic art, or machinic creation, we will not be referring to processes that are started by humans, where machinic systems are used by human artists and eventually endowed with varying degrees of autonomy in the development artworks (Galanter 2006). We will not be discussing anthropocentric procedural, algorithmic, or computational art, or what we may call *generative* art or interactive art (Carvalhais 2016). In short, we will not be discussing art produced *by* humans or *for* humans. We will be discussing the possibility of posthuman art (Bogost 2012). We will do this by focusing on the possibility of

---

[1] If any scepticism regarding human-level or above artificial intelligence, artificial creativity, or artificial consciousness does not allow one to fully follow this paper's arguments, please try to replace all occurrences of *artificial intelligence* by *alien intelligence*. Imagine replacing the idea of a created computational intelligence by the idea of an evolved organic intelligence that is nevertheless fundamentally different from humans or any earth-evolved intelligence. Keep in mind how Nicholas Rescher argued that alien life forms are perhaps "so alien that their science and technology is incomprehensible to us; we could never understand it as intelligence." (in Bogost 2012)

artistic creations developed by artificial minds and by those artificial minds' own volition. Naturally, from a human point of view these will be procedural, algorithmic, and computational, because they will be developed by computational systems. They will be generative, because the systems will be fully autonomous from any human artist. And they may be interactive, regardless of whether the interactions to be developed involve humans or other self-aware systems. They will not, however, be a result of a human deferring some—or even most—of the decisions to a machine, but the sole result of machinic intentions and decisions.

The problem of defining art

Dwelling upon the problem of defining *what art is* reveals a long list of putative definitions and an even longer list of rebuttals. From the representation theories of art[2]

> x represents y (where y ranges over a domain comprised of objects, persons, events and actions) if and only if (1) a sender intends x (e.g., a picture) to stand for y (e.g., a haystack) and (2) the audience realizes that x is intended to stand for y. (Carroll 1999, 25)

to the emergence of aesthetics as a discipline from the eighteen-century onwards

> x is an artwork if and only if (1) x is produced with the intention that it possess a certain capacity, namely (2) the capacity of affording aesthetic experience. (Carroll 1999, 162)

including expressionist theories of art

> x is a work of art if and only if x is (1) an intended (2) transmission to an audience (3) of the self-same (type-identical) (4) individualized (5) feeling state (emotion) (6) that the artist experienced (himself/herself) (7) and clarified (8) by means of lines, shapes, colors, sounds, actions and/or words. (Carroll 1999, 65)

and formalist theories of art

> x is a work of art if and only if x is designed primarily in order to possess and to exhibit significant form. (Carroll 1999, 115)

Each of these and other formulations of related theories were presented at one point as very promising definitions, but they all fell under the weight of art's capacity to reinvent itself. They all seem to agree, however, in proposing the existence of an intention that a creator somehow manifests through a work of art—the x—to a given audience

---

[2] For exemplar formulations of the different theories of art, and for the sake of consistency and comparability, we will use those proposed by Nöel Carroll (1999).

that has some ability to understand it. It is often difficult to grasp, however, what makes this x—i.e., the object, action or proposal *in itself*—entitled to the special status of being considered *a work of art* if we ignore the context of its creation and presentation. Andy Warhol's *Brillo Box (Soap Pads)* (1964) is a classic example of an artwork whose physical manifestation has negligible objective differences that we can use to differentiate it from the mundane object it was inspired on. Arthur C. Danto uses this example extensively when defining his concepts of *artworld*

> To see something as art requires something the eye cannot decry—an atmosphere of artistic theory, a knowledge of the history of art: an artworld. (Danto 1964, 580)

and of works of art as *embodied meanings*

> The artwork is a material object, some of whose properties belong to the meaning, and some of which do not. What the viewer must do is interpret the meaning-bearing properties in such a way as to grasp the intended meaning they embody. (Danto 2013, 38)

But what is this artworld that we must know before seeing something as art? How does the audience know they should strive to interpret meaning in a material object? These and other questions led George Dickie (1969) to expand on Danto's concept of the artworld and to propose an institutional theory of art that states that

> x is an artwork in the classificatory sense if and only if (1) x is an artifact (2) upon which someone acting on behalf of a certain institution (the artworld) confers the status of being a candidate for appreciation. (Carroll 1999, 227)

This theory became very successful by refocusing the attention away from the artefact and towards the sociocultural context of its presentation. It does not, however, provide any criteria to this status of being a candidate for appreciation. In the case of a given machinic artefact, the mere recognition of it as art by any agent or institution of the artworld would instantly make it art and the reasons presented for that recognition—whichever they would be—would become a way to understand the work: problem solved. The theory is unable to help us in excluding the possibility of misunderstanding the proposal or even the likelihood of considering as a candidate for appreciation something that the machine did not intend as such.

Jerrold Levinson (1979) took the institutional theory of art as an inspiration, but aimed to develop a new theory that was more focused on the intentions of the individual who created or presented the artefact or action. The historical definition of art thus states that

> x is an artwork if and only if x is an object of which it is true that some person or persons (1) who have a proprie-

tary right over x (2) nonpassingly intend (or intended) x for regard-as-a-work-of-art—i.e., for regard in any way (or ways) in which objects already in the extension of "artwork" are or were correctly or standardly regarded. (Carroll 1999, 243)

This theory might be viewed as entailing an anthropocentric view of art, since it presupposes that the artefact, action, or proposal was conceived to be regarded in a way similar to those of previous—and strictly human-conceived—artistic proposals. But, in fact, it has an elegant way of integrating the idea that art is an open, evolving concept that we can disentangle by looking for Wittgensteinian family resemblances. Also, it captures the intention of the creator that was present in previous theories but absent from the institutional theory of art. Let us assume that we somehow know that our machine has the explicit intention of creating something to be regarded as an artwork.[3] How would we try to understand it? What do we do when we regard something as art? What do we do when we regard something as not-art?

One thing seems to be certain: regarding something as art or as not-art is usually not a mere classification, but most of the times an evaluation (i.e., a judgement of value). And this evaluation comes from the experience we have of that artwork (Dewey 1934), from the relation we establish and the meaning we extract from it: we do not merely *see/hear/touch/etc.* art, but we *understand/feel* art.

As we saw, in most definitions of art we have, even if solely implied, a creator with intent and an audience capable of understanding meaning or significance. This relation is mediated by an artwork, but whilst this artwork somehow carries meaning or significance, those are not intrinsic to it, i.e., they do not constitute *objective* characteristics. Any artwork would simply cease to *be* an artwork without the existence of subjects capable of understanding and valuing it as such. Just as a gesture is never merely a physical movement, but it always conveys—or bears, as implied by its etymology—an intention from one agent to another. Stripped of this embedded meaning, a gesture can be objectively described as a movement, but in doing so we miss all of its intent, purpose and usefulness.[4]

This meaning, however, is usually not explicit. The artwork does not describe or define meaning as, e.g. a mathematical expression would. Just as with a gesture, the understanding of meaning can be altered by our attention, simultaneous stimuli, momentary feelings or our comprehension of the context. Our perception of emotion in music, e.g., has been shown to be greatly dependent on cross-modal interactions and to even be highly susceptible to

knowledge about the authors' emotional context when writing it (Margulis, 2017). If we were aiming to approach music *objectively*, then this instability of reception could be regarded as yet another proof of how unreliable and biased our perception and cognition are. But it does not make sense to approach art *objectively*. The fruition of art is always about the *subjective* experience of the artwork and anything that contributes to enrich our experience of an artwork is not an obstacle, but a catalyser. And we all share a long history of suspending our disbeliefs for the sake of our experiences of art. Furthermore, this suspension of disbelief—or even the creation of new beliefs (Gottschall 2012; Zacks 2015)—is often surprisingly effortless, and meaning spontaneously emerges in our minds as more than the sum of its explicit parts, just as when we understand a beautiful gesture or a compelling metaphor. Anything that brings us closer to the ability to understand intentions behind an artwork, a gesture or a metaphor—even if based on fictitious cues, as long as we do not know or do not actively choose to focus on the fact that they are false (Zacks 2015)—deepens our experience. But how can art do this to us? And how can we approach art in a meaningful way despite this apparent volatility?

Approaching art subjectively

We propose that the main strategy we use when trying to understand an artistic proposal is the subjective inference of the *subjective choices* made by its author. *Choices* because we assume that the artist had various options and freely selected the ones they presented. *Subjective*, on one hand, because we always assume the existence of a subject that made the choices: the author might be dead (in the Barthesian sense) or unknown—or we might misinterpret them entirely—, but we always infer their existence. *Subjective*, on the other hand, because we tend to conjecture *intentions* behind the choices. In this sense, the appreciation of art is very close to the moral appreciation of other people's actions and gestures: both assume the existence of an uncoerced choice and both are based on the reading of intentionality behind actions. As we don't have direct access to other minds, we are constricted to a second or third person view over their behaviours as our gateway to know their intentions. Yet we seem somehow able to avoid the reductive aspects of behaviourism.

We do that by reading behaviours alien to us with the superposition of our own experiences and feelings in order to build a plausible model of the intentions behind them. Artworks, in the sense that we take them to be the outcome of subjective choices, are thus interpreted as direct or indirect results of the behaviours of other human beings. In the deciphering of these artworks, as in moral discernment and gesture interpretation, we make extensive use of an astonishing human skill: *empathy*, i.e., the capability of putting ourselves in the other's shoes, of thinking "what would I do/feel/think if was seeing/experiencing/acting on the world from that perspective". For that, we make extensive use of our imagination and, according to recent proposals

---

[3] We should bear in mind, however, that it would be perfectly acceptable for our machine to conceive a completely new and unwarranted meaning to the word *art* before producing artworks: our problem would just be a lot bigger.

[4] For a good example of the problems raised by an objective approach to art, see Danto's take on the restoration of the Sistine Chapel (2013).

in neuroscience, a particular kind of neurones that fire both when we act and when we observe the same action performed by others: the mirror neurones or, as V.S. Ramachandran puts it, *"the neurons that shaped civilisation"* (2011, 117). Contrary to what can be suggested by the word *mirror*, however, the purpose of these neurones is not merely to mimic other peoples' actions. Instead, their behaviour is influenced by our predictions of what is likely to happen, according to the inferred intentions of others, and fire differently according to that interpretation:

If mirror neurons do, in fact, signal intentions, how do they do it? One possibility is that the response of these neurons is determined by the chain of motor activities that could be expected to happen in a particular context […]. For example, when a person picks up a cup with the intention of drinking, the next expected actions would be to bring the cup to the mouth and then to drink some coffee. However, if the intention is to clean up, the expected action might be to carry the cup over to the sink. According to this idea, mirror neurons that respond to different intentions are responding to the action that is happening plus the sequence of actions that is most likely to follow, given the context. (Goldstein 2013, 168)

Being associated with actions, mirror neurones are highly related to our internal mapping of our own body and we use this somatic resonance extensively, along with our experience of the limits of our own body and of the forces exerted over it by the environment, when appreciating art. We can feel weightless when watching Nureyev, overwhelmed by the physical effort of Martha Argerich playing Liszt's *B minor sonata* or inhumanely balanced when staring at *The Dancing Shiva*.

One day around the turn of the twentieth century, an elderly firangi ("foreigner" or "white" in Hindi) gentleman was observed gazing at the Nataraja in awe. To the amazement of the museum guards and patrons, he went into a sort of trance and proceeded to mimic the dance postures. A crowd gathered around, but the gentleman seemed oblivious until the curator finally showed up to see what was going on. He almost had the poor man arrested until he realized the European was none other than the world-famous sculptor Auguste Rodin. Rodin was moved to tears by The Dancing Shiva. In his writings he referred to it as one of the greatest works of art ever created by the human mind. (Ramachandran 2011, 238)

Not all art is, however, as directly relatable to the human body as the traditional performing arts. Nonetheless, reading a purpose behind actions depends on our ability to understand or imagine the possible motives other human beings had to act the way they did and we inevitably bring our own experiences, our own body, and our own culture to the table when judging an intention. We can find some object beautiful or interesting, but we call it art solely if we

can infer that someone had the intention of making it—or presenting it—as art. That is why we tend to be particularly careful in the way we experience something we find displayed at a museum, trying hard to unlock an intention behind it. In an example that travelled the world in May 2016, two students visiting the San Francisco Museum of Modern Art left a pair of "glasses on the floor below an official-looking piece of paper to see how it would be received by gallery-goers." (Hunt 2016) Pictures of people admiring and photographing this "artwork" travelled the world, through both social and conventional media, in most cases framed as evidence of just how gullible gallery-goers are or as how meaningless and decadent art has become. But it could just as easily be seen as evidence of how much art—in the classificatory sense—has the power to make us question and seek for an understanding of intentionality behind—or meaning embedded in—any object, action or proposal.[5]

Sometimes we can't understand a valid intention for presenting it as art, as sometimes we can't understand why that other person cut us off in traffic: in both cases, aesthetic and moral, we can easily feel angry or defrauded. And it seems clear that the closer our own experience is to the experience of a given artist, the easier it will be for us to understand (i.e., relate to their choices) and judge (i.e., conjure other possibilities and judge them against their choices) their proposal. The fact that it makes no sense to judge, e.g., John Cage's *Imaginary Landscape No. 4* (1952) for the content of a random radio program that happens to be heard during a given performance is only obvious to someone who understands where to find the subjective choices of the composer, despite the fact that its appearance is a direct consequence of these choices. But for those of us who understand that fact—either because we read it on the program notes or because we can interpret the cues from the context—, it becomes very clear that different performances of this piece, despite their superficial differences, are merely instantiations of the same piece. And we are capable of understanding that fact precisely because of our knowledge of where to find the subjective choices of the author—i.e., the things that don't change from one instantiation to another, the choices that govern the unpredictable elements. That is also the reason why all of the outcomes of a generative artwork governed by unaltered human-chosen algorithms, pseudo-random sources, or fitness functions should be regarded as different instantiations of the same artwork. If the subjective choices are the same, the artwork *is* the same, regardless of the range of different results that may emerge from the same process. For an artwork to be attributed to a machine, it *has* to be—even if only partially—the result of subjective choices made *by that particular machine*. As a human creator, that machine has to *be* a subject, it has to have intentions and it has to be able to make uncoerced choices

---

5 The fact that we sometimes feel compelled to avoid expressing our honest opinions about artworks is an unrelated—albeit interesting—matter.

in order to make art. It has to have a first-person perspective, i.e., it has to be sensible to ask "What is it like to be that machine?". And in order to understand an artwork made by that machine, we must be able to understand its subjective choices. For that, we need to be able to empathise with the machine, we need to be able to provide some satisfactory answer to the aforementioned question.

The problem is that we have trouble empathising with agents that we cannot relate to, with subjective experiences and actions that we have trouble understanding, perhaps the ones our neurones have more trouble mirroring. As Thomas Nagel puts it, when referring to our ability to know what is it like to be a bat:

Our own experience provides the basic material for our imagination, whose range is therefore limited. It will not help to try to imagine that one has webbing on one's arms, which enables one to fly around at dusk and dawn catching insects in one's mouth; that one has very poor vision, and perceives the surrounding world by a system of reflected high-frequency sound signals; and that one spends the day hanging upside down by one's feet in an attic. In so far as I can imagine this (which is not very far), it tells me only what it would be like for me to behave as a bat behaves. But this is not the question. I want to know what it is like for a bat to be a bat. Yet if I try to imagine this, I am restricted to the resources of my own mind, and those resources are inadequate to the task. I cannot perform it either by imagining additions to my present experience, or by imagining segments gradually subtracted from it, or by imagining some combination of additions, subtractions, and modifications. (Nagel 1974, 439)

We can empathise with non-human characters in human-made art *precisely* because they were conceived from the only subjective perspective that is accessible to us: as "what is it like for a human to behave as x". In a word, they were anthropomorphised:

The risk of falling into anthropocentrism is strong. Indeed, I'll take things farther: anthropocentrism is unavoidable, at least for us humans. The same is true of any unit (for the bats, chiropteracentrism is the problem). The subjective nature of experience makes the unit operation of one of its perceptions amount always to a caricature in which the one is drawn in the distorted impression of the other. This is true not only of the encounter itself but also of any account of the encounter, which only further distances the one from the other by virtue of the introduction of additional layers of mediation. (Bogost 2012, 64-5)

Cultural differences between humans alone can account for severe difficulties in the development of empathy and in understanding artistic manifestations, as the history of European colonialism easily shows. The perimeter of our own particular experience of the world can also impose

limits to our personal understanding of art: if, e.g., the only way I can envision myself exhibiting a given artistic proposal is by despising and deliberately trying to make fun of the audience, I will likely get offended by that proposal. It is as if any work of art could "be understood only by someone who has himself already had the thoughts that are expressed in it—or at least similar thoughts." (Wittgenstein 2001, 3).[6] And yet we seem to set no boundaries to our eagerness to read anthropomorphised creative intentions everywhere. Marvelling at the diversity and sublime might of Nature whilst reading plausible intentions of a human-like (albeit infinitely more powerful) designer, e.g., might explain our collective drive to believe in the actions of divine creatures. In the words of Ramachandran:

But we are so deeply hardwired for imputing things such as motive, intent, and culpability to the actions of others that we often overextend our social emotions to nonhuman, nonsocial objects, or situations. You can get "angry" with the tree branch that fell on you, or even with the freeways or the stock market. It is worth noting that this is one of the major roots of religion: We tend to imbue nature itself with human-like motives, desire, and will, and hence we feel compelled to sup plicate, pray to, bargain with, and look for reasons why God or karma or what have you has seen fit to punish us (individually or collectively) with natural disasters or other hardships. This persistent drive reveals just how much the self needs to feel part of a social environment that it can interact with and understand on its own terms. (Ramachandran 2011, 252)

Popular culture, e.g., has several reports of "a few animals [that] are prodigious producers of 'art'" (Goldman 2014). Of these, a small Japanese pufferfish that builds intricate circular-shaped sand structures on the seabed has been called by David Attenborough "nature's greatest artist" (2014). It is actually quite astonishing how such a small fish can build such beautiful large structures, even using shells to decorate some of its elements. Except that it is not decoration or an aimless subjective manifestation, but a strategy to attract the female pufferfish. Something that has, since Darwin, served as an implausible explanation for the prevalence of art in human cultures. The male pufferfish, once that goal is fulfilled, pays no further attention to the structure, which defies our own expectations of an artistic motivation. Even in a scientific report published by *Nature*, we can find scientists writing things such as "strangely enough, the males never reuse the nest" (Kawase, Okata & Ito 2013). But why is it strange? Because we expect a human artist to see their creations as ends and not as means?

If animals did produce art, that art would be a manifestation of their subjective choices. There is, to the best of our

---

[6] Wittgenstein refers to the contents of his *Tractatus Logico-Philosophicus*, that he introduces with this sentence, not to any work of art.

knowledge, no evidence that any animal does that. That does not mean that they do not produce art, it only means that we are unaware of that endeavour. As Ludwig Wittgenstein puts it, "if a lion could talk, we could not understand him" (2009, 235). Similarly, if a lion produced art, we could not understand it. Instead, what we find is evidence that we attribute anthropomorphised intentions to the animals that produce objects that we find aesthetically pleasant: i.e., we ask "what could motivate a fellow human to make those aesthetic choices?" when regarding some animal-built artefact or structure. We usually remain oblivious to the fact that the animals that produce human-comparable "artworks", such as paintings, often have no choice at all:

> Our results suggest that painting does not improve the welfare of elephants and that its main benefit is the aesthetic appeal of these paintings to the public and their subsequent sale of which a percentage of funds might be donated toward conservation of the species. (English, Kaplan & Rogers 2014, 14)

We make sense of the world solely from our own perspective—however wide that perspective may be—, and art is no exception to that. Given this track record of misunderstandings, we believe it is very unlikely that we will ever be able to understand machinic art in any meaningful way, i.e., in a way that is plausible from the perspective of the machine.

An example: what is it like *to see*?

If we are able to see, we simply cannot understand what it is like to be born blind, what is it like to experience the world with eyes embedded in a different part of our body (such as our legs), or what is it like to see a different part of the electromagnetic spectrum. For us, *to see* is always a shortening of *to see like me*. And we often take *computer vision* as a shortening of *making a computer see like me*. This is, as we know, far more difficult than some recent breakthroughs might suggest at first sight. Understanding machinic visual art, on the other hand, would require us to *see like a computer*. This, however, is not as simple as looking at a picture or video taken by a webcam.

Bertrand Russell, in the introduction to his *The Problems of Philosophy,* makes a distinction between the view of the practical man and the philosopher, on one hand, and that of the painter:

> […] the painter has to unlearn the habit of thinking that things seem to have the colour which common sense says they 'really' have, and to learn the habit of seeing things as they appear. Here we have already the beginning of one of the distinctions that cause most trouble in philosophy—the distinction between 'appearance' and 'reality', between what things seem to be and what they are. The painter wants to know what things seem to be, the practical man and the philosopher want to know

what they are; but the philosopher's wish to know this is stronger than the practical man's, and is more troubled by knowledge as to the difficulties of answering the question. (Russell 2001, 2)

This is reminiscent of the Lockean distinction between primary qualities—properties which an object possesses independent of the subject—and secondary qualities—caused by the interaction of a subjects' perception with the primary qualities of an object (Uzgalis 2017). If we want to know how things *really* are, we know that we should strive to get as close as possible to the former. But art, as Russell points out, is not about how things *really* are, but instead about how we *experience* things. If anything, Russell's description is incomplete: the painter wants to know what things seem to be *from their perspective*, i.e., as they themselves experience it. It is not the case that the painters are not concerned with the reality—often reduced to the reality of how things *really are—*, they simply are focused on the reality of how things *appear to them*. As Nagel puts it:

> Very little work has been done on the basic question (from which mention of the brain can be entirely omitted) whether any sense can be made of experiences' having an objective character at all. <u>Does it make sense, in other words, to ask what my experiences are really like, as opposed to how they appear to me?</u> We cannot genuinely understand the hypothesis that their nature is captured in a physical description unless we understand the more fundamental idea that they have an objective nature (or that objective processes can have a subjective nature). (Nagel 1974, 448)

Learning this, Russell's painter can then proceed to manifest, both consciously and unconsciously, their vision of how things appear to them in a holistic way: including their qualia, their underlying emotions and feelings, their views over their motifs, their sense of structure, their painting techniques, the spirit of their epoch, etc. Representational art might favour the manifestation of qualia, expressionist art might prefer to reinforce emotions, formalist art might focus on the appreciation of structure, but they all manifest a holistic subjective view of the painter to a given audience. Perspective, as John Berger puts it, is a very good example of this:

> The convention of perspective, which is unique to European art and which was first established in the early Renaissance, centres everything on the eye of the beholder. It is like a beam from a lighthouse—only instead of light travelling outwards, appearances travel in. The conventions called those appearances reality. Perspective makes the single eye the centre of the visible world. Everything converges on to the eye as to the vanishing point of infinity. (Berger 1972, 16)

The invention of photography in the 19th century, with its more *objective* view over the visual world, changed

things dramatically. Nonetheless, photography had to struggle for its status as an artistic medium, precisely because it was for a long time unclear how it could be used to manifest a subjective perspective.

> The camera isolated momentary appearances and in so doing destroyed the idea that images were timeless. Or, to put it in another way, the camera showed that the notion of time passing was inseparable from the experience of the visual (except in paintings). What you saw depended upon where you were and when. What you saw was relative to your position in time and space. […] Every drawing or painting that used perspective proposed to the spectator that he was the unique centre of the world. The camera—and more particularly the movie camera—demonstrated that there was no centre. (Berger 1972, 18)

This contradicts the common-sense idea that painting evolved to slowly approach the "realism" of photography. In fact, and from an artistic perspective, it is easy to see that the exact opposite is actually closer to the truth. Whilst the invention of photography had a significant impact on the subsequent evolution of painting, that movement was not towards the realism of the more objective view we suddenly had easier access to, but actually away from it and towards a greater "subjective realism". Photography still had to evolve, as a medium, to approach the way painting could already be used to convey a first-person perspective over something, the way that painting embodied subjective choices, the way that painting approached three-dimensionality, the way that painting approached the time from the perspective of consciousness, as opposed to the time interval determined by the exposure of film. As Danto puts it:

> With a film speed of ASA 160 and shutter speeds of one-sixtieth of a second we could now capture the face appearing in ways which the eye never sees—'between expressions,' as it were. That is why we reject as not 'really me' many of the images on a contact sheet, which don't look like what we see in the mirror… The still shows 'optical truth' but it does not correspond to perceptual truth, namely how we see the world stereoptically. (Danto 2013, 106)

We would hardly call art to a painting that aimed towards an *optical truth*, except perhaps in the common-sense idea that art can be synonymous with a highly developed skill. We had to wait until photographers learned how to manifest their subjective choices through the use of the camera and until audiences learned how to empathise with the human behind the lens to understand photography as an art form. What if behind the camera—or the webcam—is not a human, but a machine capable of making subjective choices? Will we ever learn how to empathise with the subjective views of a being that has such a radically different experience of the world?

Conclusion

The problem is not whether machines will or will not develop a sense of self that leads to an eagerness to manifest their own subjective experiences of the world. The problem is that if—or *when*—they do, they will have such a different experience of the world that we will likely be completely unable to relate to it from our subjective perspective. Our subjective human experience stems, amongst many other things, from being born and slowly educated within a society of fellow human beings, from fighting the inevitability of our own death, from saving memories based on our own insubordinate feelings, from the lonely curiosity of our own mind, from the omnipresence of the needs and quirks of our biological body and from the way it dictates the space and time scales we can grasp.

It may very well happen that we understand some actions or artefacts created by machines of their own volition as art, but in doing so we would most likely be anthropomorphising the machine's intentions and thus missing out on their machinic perspective. Whilst we can have interpretations of a human-made artwork that differ from those of the author—art does not describe with observer-independent purposes—, these interpretations infer a human author and, if well-informed, can be reasonable even for the original author. With a non-human author—and assuming we get to the point of correctly identifying the artefact as an artistic manifestation—, our anthropomorphised interpretation will likely seem implausibly alien for a machinic perspective.

On the other hand, we may of course envision building a biological machine that would be born and afraid of dying just like us, a machine that would be emphatic and social just like us, that would have a body, sensorial apparatus, and mirror neurones just like our own. In face of that machine, we would certainly be much more likely to understand their actions or artefacts as art. But maybe we should more accurately call that machine a *human being*.

Acknowledgements

References

Anderson, Alan R., ed. *Minds and Machines*. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Ariza, C. (2009). The Interrogator as Critic: The Turing Test and the Evaluation of Generative Music Systems. *Computer Music Journal, 33*(2), 48-70. doi:doi:10.1162/comj.2009.33.2.48

Attenborough, D. "Life Story, episode 5: Courtship." *BBC Two.* November 20, 2014.
http://www.bbc.co.uk/programmes/b04q1rwy

Bateson, G. (1979). *Mind and Nature: A Necessary Unity.* New York: E. P. Dutton.

Berger, J. *Ways of Seeing.* London: BBC and Penguin Books, 1972.

Bogost, I. (2006). *Unit Operations: An Approach To Videogame Criticism.* Cambridge, Massachusetts: The MIT Press.

Boden, M. A. (2004). *The Creative Mind: Myths and Mechanisms.* (Second ed.). London: Routledge.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies.* Oxford: Oxford University Press.

Carroll, N. *Philosophy of Art: A Contemporary Introduction.* London: Routledge, 1999.

Carvalhais, M. (2016). *Artificial Aesthetics: Creative Practices in Computational Art and Design.* Porto: U.Porto Edições.

Cope, D. (2005). *Computer Models of Musical Creativity.* Cambridge, Massachusetts: The MIT Press.

Damásio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain.* New York: Pantheon Books.

Damásio, A. (2017). *A Estranha Ordem das Coisas* (L. Oliveira Santos, Trans.). Lisboa: Temas & Debates.

Danto, A. C. "The Artworld." *The Journal of Philosophy* 61, no. 19 (1964): 571-84.

Danto, A. C. *What Art Is.* New Haven, CT: Yale University Press, 2013.

DeLanda, M. (1991). *War in the Age of Intelligent Machines.* New York: Zone Books.

Dennett, D. C. (1991). *Consciousness Explained.* London: Penguin Science.

Deutsch, D. (2011). *The Beginning of Infinity: Explanations that Transform the World.* London: Allen Lane.

Dewey, J. *Art as Experience.* New York, NY: Perigee Books, 1934.

Dickie, G. "Defining Art." *American Philosophical Quarterly* 6, no. 3 (1969): 253-56.

DiSalvo, D. (2011). *What Makes Your Brain Happy and Why You Should Do the Opposite.* Amherst, NY: Prometheus.

Eagleman, D. M. (2011). *Incognito: The Secret Lives of the Brain.* New York: Pantheon Books.

English, M., Kaplan, G., and Rogers, L. J. "Is painting by elephants in zoos as enriching as we are led to believe?" *PeerJ* 2:e471 (2014).

Galanter, P. (2006). Generative Art and Rules-Based Art. *Vague Terrain.* Retrieved from http://philipgalanter.com/downloads/vague_terrain_2006.pdf

Gazzaniga, M. S. (2011). *Who's in Charge?: Free Will and the Science of the Brain.* New York: Ecco.

Goldman, J. G. "Creativity: The Weird and Wonderful Art of animals." *BBC Future.* July 24, 2014.
http://www.bbc.com/future/story/20140723-are-we-the-only-creative-species

Goldstein, E. B. *Sensation and perception.* Ninth ed. Belmont, CA: Wadsworth, Cengage Learning, 2013.

Gottschall, J. (2012). *The Storytelling Animal, How Stories Make Us Human.* New York, NY: Houghton Mifflin Harcourt, 2012.

Hayles, N. K. (2005). *My Mother Was a Computer: Digital Subjects and Literary Texts.* Chicago, IL: The University of Chicago Press.

Hofstadter, D. (2007). *I Am A Strange Loop.* Cambridge, Massachusetts: Basic Books.

Humphrey, N. (2011). *Soul Dust: The Magic of Consciousness* Princeton, New Jersey: Princeton University Press.

Hunt, E. *"Pair of glasses left on US gallery floor mistaken for art." The Guardian.* May 27, 2016.
https://www.theguardian.com/us-news/2016/may/27/pair-of-glasses-left-on-us-gallery-floor-mistaken-for-art

Jefferson, G. "The Mind of Mechanical Man." *British Medical Journal* 1, no. 4616 (1949): 1105–10.

Kawase, H, Okata, Y., and Ito, K.. "Role of Huge Geometric Circular Structures in the Reproduction of a Marine Pufferfish." *(Nature) Scientific Reports* 3, no. 1 (2013).

Kelly, K. (2016). *The Inevitable.* New York: Penguin Books.

Levinson, J. "Defining Art Historically." *British Journal of Aesthetics* 19, no. 3 (1979): 21-33.

Margulis, E. H. "Music is not for ears." *Aeon Essays.* November 2, 2017. https://aeon.co/essays/music-is-in-your-brain-and-your-body-and-your-life

Metzinger, T. (2009). *The Ego Tunnel: The Science of the Mind and the Myth of the Self.* New York: Basic Books.

Morowitz, H. J. (2002). *The Emergence of Everything: How the World Became Complex.* New York: Oxford University Press.

Nagel, T. (1974). What is it Like to be a Bat? *The Philosophical Review, 83*(4), 435-450.

Oliveira, A. (2017). *The Digital Mind: How Science is Redefining Humanity.* Cambridge, Massachusetts: The MIT Press.

Penrose, R. *The Emperor's New Mind: Concerning Computers, Minds and The Laws of Physics.* Oxford: Oxford University Press, 1989.

Pinker, S. (1999). *How the Mind Works.* London: Penguin Books.

Ramachandran, V. S. (2011). *The Tell-Tale Brain: A Neuroscientist's Quest for What Makes Us Human.* New York: W. W. Norton & Company.

Russell, B. *The Problems of Philosophy.* Oxford: Oxford University Press, 1912. 2001.

Searle, J. R. (1981). Minds, Brains, and Programs. In D. Hofstadter & D. C. Dennet (Eds.), *The Mind's I: Fantasies and Reflections on Self and Soul* (pp. 353-373). New York: Basic Books. (Reprinted from: 2000).

Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York: Alfred A. Knopf.

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, 59*, 433-460.

Turing, A. M. (1996). Intelligent Machinery, A Heretical theory. *Philosophia Mathematica, 4*, 256-260.

Uzgalis, W. "John Locke." In *The Stanford Encyclopedia of Philosophy (Fall 2017 Edition),* edited by Edward N. Zalta. April 24, 2017.
https://plato.stanford.edu/archives/fall2017/entries/locke/

Wagner, S.-C. (2017). *Poietry: Challenging Solitude and the Improbability of Communication* (D. Benedek, Trans.). Berlin: de Gruyter.

Wittgenstein, L. (1921). *Tractatus Logico-Philosophicus* (D. F. Pears & B. F. McGuinness, Trans.). Abingdon: Routledge.

Wittgenstein, L. *Philosophical Investigations*. Translated by G.E.M. Anscombe, P.M.S Hacker and J. Schulte. Revised 4th edition by P.M.S. Hacker and J. Schulte. West Sussex: Wiley-Blackwell, 1953. 2009.

Zacks, J. M. (2015). *Flicker: Your Brain on Movies*. Oxford: Oxford University Press.

Authors Biographies

Rui Penha is a composer, media artist and performer of electroacoustic music. He studies the relationship between music and its technology, developing interfaces for musical expression, sound spatialisation software, interactive installations, musical robots, autonomous improvisers and educational software. He is an invited assistant professor at the Faculty of Engineering of the University of Porto and a senior researcher at INESC TEC. http://ruipenha.pt

Miguel Carvalhais is a designer and musician. He is an assistant professor at the Faculty of Fine Arts of the University of Porto, researching computational media, interaction design, and creative practices with procedural systems. He is the author of a book on these topics, *Artificial Aesthetics*. http://carvalhais.org