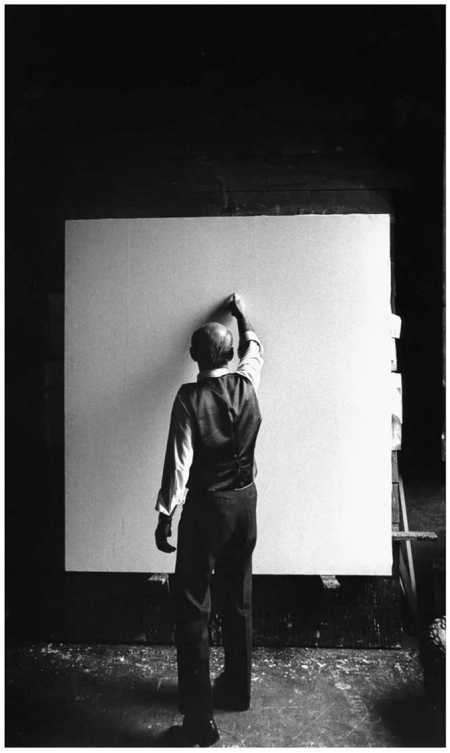
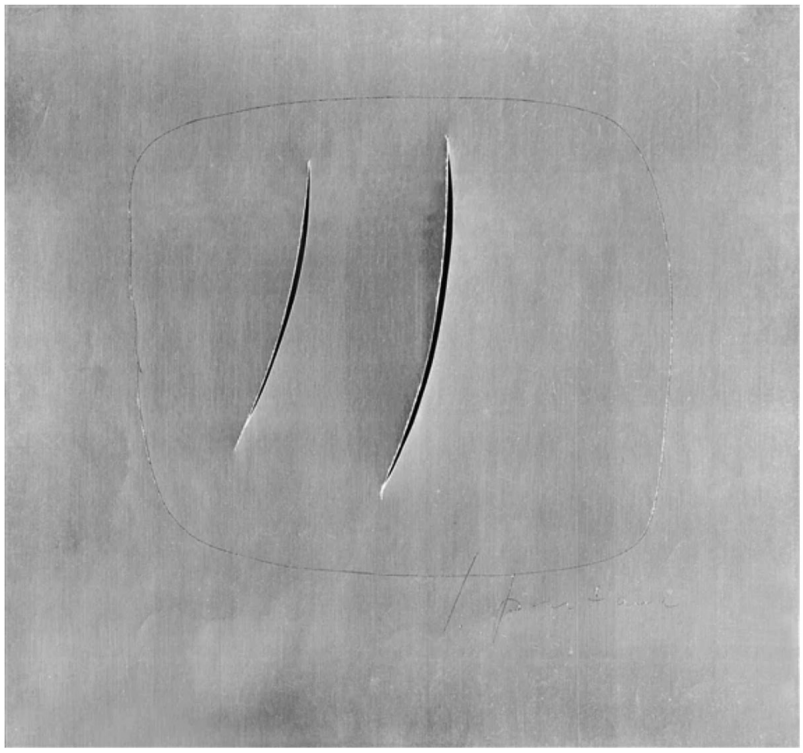
AI-aesthetics and the Anthropocentric Myth of Creativity

Emanuele Arielli and Lev Manovich





Previous page: Lucio Fontana making one of his cut paintings, 1964.
This page: Hatsune Miku performing at the 2020 Coachella Valley Music and Arts Festival.

Artificial aesthetics

Since the beginning of the 21st century, computation, data analysis, and artificial intelligence have gradually entered the aesthetic realm, being used in recommendation systems for art, music, books, and movies or in the automatic editing of images and video. AI is also increasingly used to generate new synthetic artifacts, including artworks, music, designs, and texts. For instance, in 2016 a deep-learning algorithm was trained to learn Rembrandt's style by analyzing his 346 known paintings and was then asked to generate a brand-new portrait. The result looked uncannily like a real Rembrandt painting. In the same year, researchers of the Sony Computer Science Laboratories in Paris developed a neural network called DeepBach, producing choral cantatas

in the style of J.S. Bach.[1] Since then, other music generating algorithms have been created – today YouTubers challenge viewers to take part in musical "Turing Tests" by differentiating AI-music compositions from human ones. For people with some musical training, the task seems straightforward, but this is not always the case for naïve listeners.[2] In 2019, an AI used the computing power of a new model of smartphone to finish Schubert's "Unfinished Symphony" (n. 8, 1822)[3]. However, this was accomplished with the help of a composer who did a bit of cherry-picking by selecting the best melodies generated by the AI. In the same year, Deutsche Telekom organized a team of international music and AI experts to complete Beethoven's unfinished 10th symphony and thus celebrate the 250th anniversary of his birth. The completed symphony "Beethoven X - The AI Project" premiered on October 9, 2021, in Bonn. In these examples, computers are fed with pre-existing styles and, in turn, generate variants conforming to those styles, trying to introduce some innovation. They do not generate completely new songs or styles; instead, they seem to be examples of what we might call computational *mannerism*.

When a machine paints a Rembrandt, composes a Bach sonata, or completes a Beethoven symphony, we say that this is neither original nor real *art*, but simply the complex imitation and reproduction of existing products of human culture. We face the old question concerning the nature of creativity: what kind of recombination of ideas, unusual analogies, and conceptual connections are considered the mark of originality? To whom should we attribute authorship if an artifact or image is the product of devices, algorithms, and technological extensions that generate and reinterpret an artist's or designer's intention? Since the production chain is mediated by increasingly complex intervention from third-party software (as in photo and video effects and filters or retouching algorithms), how can we determine where the creative innovation

[1] https://arxiv.org/abs/1612.01010

[2] https://www.youtube.com/watch?v=PmL31mVx0XA; https://www.youtube.com/watch?v=lv9W7qrYhbk

[3] https://www.classicfm.com/composers/schubert/unfinished-symphony-completed-by-ai/

has taken place and who its author is? According to artist Mario Klingemann, one of the pioneers in AI-art: “If you heard someone playing the piano, would you ask?: “Is the piano the artist?” No. So, same thing here. Just because it is a complicated mechanism, it doesn’t change the roles”. From this perspective, AI use in the arts would be a simple instance of *extended aesthetics*, using new, apparently more sophisticated devices under the authorial control of the human artist. An artificial system would be the artist’s and programmer’s tool, a sophisticated instrument deployed during creation. However, we are still fascinated by the idea that we could also witness the emergence of autonomous artificial creativity in the aesthetic domain, holding to the original idea of true artificial intelligence as the manifestation of autonomous and intelligent behavior.

On aesthetic “Turing Tests”, or: What do we expect from “aesthetic” machines anyway?

In 2020, a Princeton University undergraduate student used for her senior project a so-called Generative Adversarial Network (GAN) to produce traditional Chinese landscape paintings that were able to fool humans in a visual Turing Test.[4] In its original formulation, the Turing Test by Alan Turing (1912-1954) is a criterion for being able to say that an artificial system has achieved human-like intelligence. However, we would not say that the GAN developed by the Princeton student reached human-level intelligence; it is just a program sophisticated enough to generate images that appear to be man-made. This further contributes to conceptual confusion in this discussion.

On the one hand, notions such as "intelligent" or "creative" seem intuitive and straightforward, so that everyone would be able to recognize intelligent or creative behavior when they manifest it themselves. On the other hand, when

[4] https://arxiv.org/pdf/2011.05552.pdf

we try to give a working and operational definition of these notions, we see how elusive they are. This issue sets Alan Turing in opposition to Ludwig Wittgenstein (1889-1951), who believed that we need first to clarify our linguistic and conceptual habits when we want to understand what we mean by terms like "intelligence". Turing attended Wittgenstein's lectures on the philosophy of mathematics in 1939 and the latter was certainly aware of Turing's thesis about mechanical thinking. Interestingly, Wittgenstein's opinion is expressed in passages such as the following, taken from his *Philosophical Investigations* (1953):

> "Could a machine think?——Could it be in pain?—Well, is the human body to be called such a machine? It surely comes as close as possible to being such a machine. But a machine surely cannot think!—Is that an empirical statement? No. We only say of a human being and what is like one that it thinks. We also say it of dolls and no doubt of spirits too. Look at the word "to think" as a tool (Wittgenstein, 1953: pp. 359-360)."

From Wittgenstein's point of view, since words are tools, we need to ask ourselves under which condition – if any – we would use notions like "thinking" (or "intelligence" and "creativity") to describe non-human, artificial entities.

The Turing Test is a method to verify if a machine talking through a computer interface would pass as human. Therefore, the test considers *mimicry* of human behavior as an indicator for intelligence, primarily focusing only on verbal cues and dialogue generation. On one hand Turing's criterion seems reasonable: if something is not distinguishable from a human in a conversation, why not attribute intelligence to it? On the other hand, however, humans are reluctant to easily grant the mark of intelligence to non-human entities. In the past, it was thought that a machine capable of beating a Grandmaster at chess would demonstrate to be a true AI. This happened in 1997, when DeepBlue beat world champion Garry Kasparov. At that point

chess was defined as a mere combinatorial and computational game, not as a true test of intelligence; the goalpost was moved to other games like Go, considered more complex and based more on creative intuitions. However in 2016 Google's AlphaGo beat world champion Lee Sedol (b. 1983), yet we do not feel like saying that a "true" intelligence has been achieved. Or consider chatbots. According to Turing's 1950 paper[5], by the end of the century machines would be able to fool a third of people after five minutes of conversation. In 2014, 33% of judges considered chatbot "Eugene Goostman" to be human, effectively passing Turing's test (one should note here that Goostman was programmed to simulate the volubility and the quirkiness of a 13-old teenager from Odessa, Ukraine).

Every time a technological milestone is reached, the goalpost seems to move further away. From a Wittgensteinian point of view, the reason does not lie in the fact that new technological milestones are not persuading enough to convince us that we are dealing with real intelligence. The question in fact is not at all empirical, but related to the assumptions we make in using and attributing concepts like intelligence and creativity. This leads to what has been called *Tesler's theorem*, which states that: Artificial intelligence is whatever has not been done yet (or, conversely, intelligence is whatever machines have not done yet).[6] Today, an application such as Siri may be able to conduct human-like dialogues. A text generator based on the recent GPT-3 by Open-AI – trained with a 570 GB dataset of Internet texts - can write sophisticatedly journalistic articles that are undistinguishable from human generated texts. However, precisely because we know that these are the products of sophisticated programming, we still think that there is no real intelligence, let alone attribute intentionality or consciousness to those systems. Put another way, we are not inclined to use the word "intelligence" in such a case; we commonly use it when referring to persons and, as

[5] Alan M. Turing, "Computing Machinery and Intelligence." Mind, 1950, 59, p. 433-460.

[6] The author of this definition is Larry Tesler, a well-known computer scientist who worked at Xerox PARC, Apple, and Amazon.

Wittgenstein said, words are *tools* with specific usage we are accustomed to. Therefore, a further corollary of Tesler's theorem is that every use of the term "AI", in contexts such as facial recognition, spam filters, computer vision, speech generation, and so on, is by definition *not* AI, but technology that makes use of complex optimization algorithms. It is just called “AI” for marketing reasons.

If the attribution of intelligence is a horizon line that can never be reached, one may wonder if there are human skills laying beyond that line at all: every time machines “solve” a specific human skill, this skill ceases to be real intelligence, turning out to be more mechanical than it appeared. This may have consequences on our understanding of human intelligence itself.

Arts and come into play here. The encounter between AI and aesthetics is crucial because art is considered a quintessentially human domain and its intractability and complexity has long appeared insusceptible to algorithmic reduction. Many people consider art, aesthetics, and creativity to be the pinnacle of human abilities; they are therefore seen as the last barricade against the advances of AI, lying further away from what technological progress can reproduce. If we stay with the traditional definition of the Turing Test, in the aesthetic domain this would boil down to the possibility to produce an artifact (be it a text, a dialogue, or a work of art) that is able to fool a human. But why should human art likeness be taken as a benchmark? What about innovative, beautiful, or compelling designs or art forms that clearly appear *non-human*? A Turing Test whose goal is to fool an observer would be, in this case, unsuitable.

Therefore, we may wish to revise the aim of a Turing Test beyond the simple “imitation game” it is originally based on and define its purposes differently. For example, we could say that a machine passes such a test if any of these conditions are met:

1) Achieves *superior* human performance (that is, produces something that is ranked higher in beauty, pleasantness, "amazingness," etc.), without regard to similarity of human cultural behavior.
2) Manifests the ability to be *creative*, that is, to generate novelty.
3) Shows *autonomous* behavior, in which the machine seems able to produce something unexpected, distant from the programmers' initial parameters and inputs.

A notorious example of superior performance (1) in AI is programs beating humans in games like chess or Go. But even in aesthetics, the ability to produce something that is judged to be superior to humans is not new: as early as 1966, an algorithm generated Mondrian paintings that were judged by the public to be aesthetically more pleasing than the actual Mondrian canvases.[7] This could make us think of a scenario in which artificial systems will produce superior music, better books, more compelling screenplays, not necessarily from the perspective of an art critic, but simply from that of the cultural industry: i.e. systems whose artifacts enjoy great public and commercial success. Taking the cost/revenue ratio into account, algorithms generating tunes or lyrics (or painting in the style of Mondrian or another famous artist) would surpass human production also from a purely economic perspective, since there is no trademark protection for the musical or pictorial style of an artist.[8]

Concerning creativity, this in itself is an elusive notion and the subject of long debates in philosophy and cognitive sciences. A "creativity Turing Test" is otherwise called an *Ada Lovelace test*, according to remarks on the possibility of creative machines made by the 19th century mathematician Ada Lovelace.

[7] Noll, "M. Human or machine: a subjective comparison of Piet Mondrian's "composition with lines" (1917) and a computer-generated picture." *The Psychological Record*, 1966, 16, p. 1-10.

[8] See platforms like aiva.ai that allow generating new copyright-free music following the style of existing songs.

In a test like this we would show an artifact generated by a machine and ask the public to judge if (and to what extent) it is creative[9].

Judging creativity and novelty is partly a subjective matter, often depending on how we, as humans, *attribute* creativity to a behavior. For example, one narrow interpretation presupposes that only humans could be capable of creativity and that we can speak of creative behavior only when one is *self-conscious* and aware of what one is doing. However, we also often use this concept in a more liberal and metaphorical way when, for example, we say that "nature is creative" (for example, in bringing about a new organism or a new virus). In this case, we just apply the notion of creativity to a phenomenon that is *unexpected*, i.e. to our knowledge, it did not exist before.

From this perspective, any random and surprising process that is not easily predictable should be considered creative; it is no accident that 20th century avant-garde artists like the Dadaists experimented with stochastic processes. However, random processes by themselves are not enough to call something creative: we expect something creative to be meaningful as well, such as a novel solution to old problems or a superior way to address some task or issue.

Similar to the challenges in defining creativity, defining *autonomy* is also not easy. A machine appears to be autonomous if it shows behavior independent from its original programming – that is, again, if it behaves in ways that are unexpected and unpredictable for the observer. On one hand, there is no clear-cut criterion for autonomy: is a mono-cellular organism autonomous? What about an insect? In attributing autonomy, we have a great deal of subjectivity as well.

[9] https://arxiv.org/pdf/1410.6142v3.pdf

AI as a critical mirror on human faculties

The philosopher Ludwig Wittgenstein, who discussed with Alan Turing the possibility of mechanizing computation and thought, offered a different interpretation of his famous test. According to Wittgenstein, this is not a method to see if a machine can fool an observer and pass for a human. The test would instead show to what extent *humans can be mechanical* in their processes and behaviors. If we see things from this perspective, the development of applications that simulate human creativity would have a sobering effect. For example, a program that can generate catchy melodies or compelling screenplays would reveal how much "mechanics" are core to those processes that we otherwise consider intuitive and free. A consequence would be that, no matter how we define the goal of a Turing Test, machines passing the test would show that humans are much more mechanical than we think. As a result, creativity may be overvalued as a human faculty simply because we do not understand its workings. The fact that specific human processes appear to be more mechanical and procedural than we assume challenges the typically romantic conception of creative intuition. One should remember how the idea of pure creativity originates from an exaltation of individual autonomy that has established itself only in modernity. This was not conceivable in ancient times, where the dominant view saw people as being only able to remember (in the sense of Platonic *anamnesis*), reconstruct, and reproduce things that already existed. The artist, in this sense, was a discoverer, not a creator; art was not a domain of pure invention but of craft and skillful imitation of reality. True creativity, in the ancient and medieval sense of *creatio* (*ex-nihilo*), was the prerogative of the divine only.[10]

Historical development of art styles is considered the product of unpredictable creative leaps that we can reconstruct in retrospect, but cannot predict in advance. However, some applications of evolutionary algorithms seem to hint

[10] Tatarkiewicz, Władysław, *A History of Six Ideas: an Essay in Aesthetics*, 1980, The Hague: Martinus Nijhoff.

at a different picture. For instance, concerning visual arts, Lisi and colleagues (2020)[11] showed the possibility of predicting stylistic development in the pictorial arts by training a system to extrapolate specific evolutionary laws by analyzing large databases of images and then generating images of temporally subsequent new styles. According to the authors, the system surprisingly generated predictions that closely mirror the actual evolutions that such styles underwent in the history of visual art, highlighting the "algorithmic" character of certain stylistic developments. That means that they would not be the product of historical contingencies or spontaneous inventions by unique artists, but rather the almost necessary progression of intrinsic formal laws.[12] Such a system, moreover, would also be able to predict *future* styles of visual art. Those developments do not need to be deterministic, but would nonetheless be the product of a range of finite combinations that data analysis systems could detect and reproduce.

These examples seem to lead to the conclusion that "being creative" is a label that an observer ascribes to phenomena whose underlying processes he is unaware of. For example, when *Go* world champion Lee Sedol was beaten by AlphaGo in 2016, he claimed that the program could make incredibly creative moves, revealing how certain moves or game strategies that humans thought were creative, were actually quite predictable. During the second game of the challenge, AlphaGo made a move (n. 37th) that many commentators described as unusually creative and caught the player off-guard, allowing the computer to win. The fact that this specific move was viewed as creative by the observers lies in the fact that players and experts did not have an understanding of what AlphaGo's underlying strategy was. From the machine's point of view, in fact, that move was the product of an evaluation that followed the same optimizing processes with which the system selected every other

[11] Lisi E, Malekzadeh M, Haddadi H, Lau FD-H, Flaxman S. "Modelling and forecasting art movements with CGANs.", 2020, Royal Soc. Open Sci. 7: 191569. http://dx.doi.org/10.1098/rsos.191569.

[12] A similar idea of an internal logic of the form itself was also suggested by George Kubler's *The Shape of Time*, 1962.

move. In this respect, calling something creative is often a measure of our lack of understanding: what we know is ordinary, what we do not know is deemed extraordinary. In other words, if we think humans are creative and AIs are not, this is because we better understand how AI works, while we still do not sufficiently understand how humans work. Technological advancements often seem to make evident that allegedly extraordinary phenomena are the product of ordinary processes.[13]

No ghost, just a shell?

Suppose human creativity could be potentially replicated by mechanical processes. In that case, we would face a crossroads: either we could give up using the concept of creativity altogether, or if we hold to our common understanding of what creativity is, we could agree to apply this concept to non-human phenomena as well, as world champion Lee Sedol did when judging the performance of AlphaGo.

However, the idea that artificial creativity discloses the mechanic nature of human creativity should also be met with a bit of critical detachment, particularly if we consider the specific case of the arts. In fact, artificial reproductions of human artifacts do not follow the same processes with which humans actually produced those artifacts. Nobody thinks that Mondrian followed procedures similar to the algorithm used in 1966 that generated pseudo-Mondrian, even though the public appreciated the artificial images more than the original ones. We cannot ignore the symbolic, historical, and conceptual meanings behind the painter's stylistic innovation, nor his role within the development of painting in relation to abstraction, figurative art, expressionism, and minimalism. In other words, the algorithm did not reproduce the *cultural process* through which Piet Mondrian got to his abstract

[13] Creativity consists in “extraordinary results of ordinary processes”. R.J. Sternberg, T.I. Lubart, “Investing in creativity” American Psychologist, 1996, 51, p. 681.

paintings. Instead, the programmers imitated the final product only on a formal level. We admire Mondrian's paintings as the final expression of the artist's journey that led to their production, their cultural role within the history of painting. Without these factors, we would see his paintings just as interesting geometric patterns but with no artistic value.

Similarly, a cut canvas by Lucio Fontana would be just a canvas with a cut that a mechanical arm equipped with a knife (like those already used in robotic surgery) and guided by a program would easily reproduce. The simplicity in producing those works reveals that there is more to them than their appearance, showing the separation between aesthetic and artistic value peculiar to contemporary art. In our aesthetic evaluation of these works, we see a historical, conceptual, and symbolic dimension in the object and we attribute specific intentions to the creator beyond what we can see on the formal surface of the canvas. A bundle of symbolic meanings, affective evocations, and cultural references enriches the artifact; we are ready to do this only if we see it coming from a subject to whom we attribute full consciousness of these meaning. Conversely, we are reluctant to grant significance to what is produced by an algorithm because we see it as soulless.

Techno-animism and the Pygmalion effect

Our natural tendency to attribute intentionality to phenomena is what would allow for the recognition of a machine as intelligent or even conscious. Children do that toward toys and other objects; sometimes adults too attribute human-like agency to, for instance, plants or small animals. Many present and past cultures hold a deep animist stance toward natural events that they could not explain through a causal and physicalist explanation. In these worldviews, non-human agents richly populate reality, be it plants, animals, or meteorological or geological phenomena. How would someone coming from

the Stone Age interpret, for example, the behavior of today's automatic doors sliding open every time someone steps in front of them? He probably would think that they possess intelligence and purpose. It would be naïve to define those animistic views as simply wrong: given the lack of better explanation, models based on intentionality often have good explanatory power in describing such phenomena. For the prehistoric man or woman, that door *wants* to open and let the person pass through. Similarly, our perception of AI strongly depends on how we project and attribute agency to artificial non-human entities.

Although the predisposition to attribute a soul to non-human entities depends on our cultural background, religious sensibility, and individual beliefs, today, the dominant assumption is that only humans (and, to a lesser extent, some animals) have *real* intentionality and agency. Whenever we attribute intentionality towards other entities (a door, a toy, a virtual assistant, the weather), we say we do it always just in a *metaphorical* sense, as a kind of fictional attitude in which we behave "as if" the entity had some agency, but without really believing it. This similarly happens when we engage with a movie's or a novel's characters "as if" they were real, even knowing that they are not[14]. However, it should be noted that the boundary between the perception of real agency and a make-believe one is fluid. For example, we consider pets like cats and dogs as having real intentionality. For many, this applies to insects or bacteria too, but for some, this is not the case anymore. Others, on the contrary, project personality even to plants, while others do that only in an "as if" fashion. Individual and cultural differences determine where the line between real and fictional attribution of intentionality is drawn.

As far as technological devices are concerned, we are in the realm of a "make-believe" attitude toward them: we learn to interact with virtual assistants like Alexa talking "as if" someone would listen to us like a human. As the complexity and flexibility of these devices increases, we may begin to view

[14] K. Walton, *Mimesis and make-believe*, 1990, Harvard University Press.

them as full-fledged entities endowed with agency. If this happens, one reason will undoubtedly be the advancement of those technologies. However, another reason will also be the cultural overcoming of prejudices: today, we would still rather give intentionality to an insect than to Alexa, no matter if the latter's complexity, access to knowledge, and ability to interact with us surpasses those of a bug by measure. Moreover, maybe we should question the idea that the "as if" intentionality (applied to things, animals, and non-human entities) is merely a metaphorical derivation of "true" intentionality. The opposite may be the case: the narrow conception of true intentionality (applied only to humans) would derive from the "as if" intentionality emerging from our natural and deep inclination to attribute agency to a wide range of phenomena.[15]

In this debate, we sometimes observe two apparently opposite positions: one considers real intentionality only in humans (and some animals), the other attributes agency to non-human entities, "humanizing" them through a kind of naive animism. Both positions, however, share the same anthropomorphic view of agency and intentionality, being in one case denied and in the other granted to non-human entities. An alternative view is to develop a notion of agency for sub-personal processes, non-human entities, and mechanical phenomena. Thus, it is not a matter of humanizing what is non-human but of developing an understanding of non-human and non-anthropocentric agency. In this matter, a change in our perception of AI would also result in overcoming an anthropocentric perspective of agency and creativity. This would follow the direction already outlined by classical post-human theorizations, as in the works of Donna Haraway and Rosi Braidotti, or by Bruno Latour's proposal to "re-assemble the social" through the inclusion of human and non-human entities, encompassing not only non-human natural agents (animal or vegetal) but also artificial ones.

We could add that the question of attributing agency and intentionality seems important in certain forms of cultural production, but not necessarily in others.

[15] This is notoriously the idea championed by Daniel Dennett. See D. Dennett, *The Intentional Stance*, 1987, MIT Press.

A decorative pattern, a piece of furniture, or a car do not (always) require authorial depth; we do not need to see meanings or reason about the author's thoughts. Even a catchy song that aesthetically engages us can lead us to ignore the presence or absence of the authorial intentions behind it. Similarly, a movie can be evaluated positively for the simple fact of being engaging and entertaining by itself, without having us think about what the writer or director wanted to say. The generation of AI-art thus becomes an interesting test case to determine in which areas we feel the need for a recognizable agent behind an artifact and in which we can do without one.

On the one hand, one may think that a song could be successful only if it satisfies our need to project intentionality on its writer and performer, allowing us to instill symbolic, emotional, and personal depth in it. On the other hand, the success of virtual pop-stars in east-Asian cultures (such as Hatsune Miku and several K-pop "avatar"-bands, some of them AI-driven) reveal how the public can emotionally engage with a fictional performer, follow them in their social profiles, go to their concerts, and buy gadgets depicting them[16]. We could go as far as to say: fans do not love them *despite,* but actually *because* they are openly fake. No real person gets in the way of their need to project their desire and imagination onto the surface of a virtual avatar. Like in the myth of Pygmalion, the artist who scorned real women but fell in love with their sculptural idealization, we may find a deep satisfaction in engaging with a machine whose soullessness allow us to infuse it with our idea of a perfect soul. Therefore, a fluid transition from "as if" to "real" attribution of a soul takes place: we are interested in engaging with a virtual persona in the same fashion as we root for the character of a movie series (or of a cartoon) or obsess over the protagonist of a novel. The same goes for the artist producing those songs or stories: we stop worrying whether the creative process resulted from an algorithm or was human-made, as long as we are emotionally engaged.

[16] https://www.flyfm.audio/flycelebrity-the-future-of-k-pop-all-artificial-intelligence-ai-girl-group-eternity-drops-their-mv