# ARTIFICIAL CREATIVITY AND GENERAL INTELLIGENCE

**CATERINA MORUZZI**
Universität Konstanz
caterina.moruzzi@uni-konstanz.de

https://doi.org/10.34632/jsta.2020.9481
**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99

**ABSTRACT**

It is hard to deny that the notions of creativity and intelligence are inherently connected. But what does this correlation amount to? Is creativity a necessary desideratum of intelligence? On the other hand, does the fact of being intelligent necessarily imply being creative as well? The aim of this paper is to explore these questions and to contribute to the discussion regarding the connections between the notions of creativity and intelligence. In order to do so, I draw on the results obtained from a study on the public perceptions and attitudes in relation to the use of AI in the creative sector conducted at the University of Nottingham. Through this discussion I aim to test the hypothesis that the key features of creativity correspond to aspects that are essential for the realization of Artificial General Intelligence, e.g. flexibility, domain knowledge, and common-sense. After having illustrated the parallels between the two concepts, I contend that while creativity is a crucial component of general intelligence, the constituents needed to build an AGI may not be sufficient to design creative artificial systems. I close the paper by tentatively suggesting how the motivations behind the discontent expressed by the participants against creative AI can be explained through the uncanny valley phenomenon.

Keywords: Computational creativity; Artificial general intelligence; Human-subject study.

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

## 1. INTRODUCTION

In the last decades, machine learning systems have learnt to replicate some processes of the human brain, such as the process of induction, quite successfully. However, they struggle to reproduce other processes, such as common sense and tacit knowledge, which help us adapt with flexibility to different environments. What do artificial systems need in order to overcome these challenges? Is creative thinking a requirement to achieve the flexibility needed to successfully interact with uncertain environments and with other agents?

The aim of this paper is to contribute to the debate regarding computational creativity and to the exploration of the connections that exist between creativity and natural and artificial intelligence. After a first section where I provide an overview of the history of research on Artificial General Intelligence, in the central part of the paper I present the results of a survey conducted at the University of Nottingham in September 2019 on the attribution of creativity to Artificial Intelligence (henceforth AI).

By drawing on the replies given by participants to this study, I intend to test the hypothesis (H1) that in order to acquire creative abilities AI needs to develop aspects that are required to progress from Narrow to Artificial General Intelligence (henceforth AGI). Indeed, I contend that many of the features that systems need to develop in order to achieve general intelligence are aspects that they need to possess also to earn the attribute “creative”.

In section 3, I discuss how some of the participants’ replies confirm the hypothesis that the features that state-of-the-art AI lacks and that would instead be essential to creativity correspond to the aspects of intelligence which researchers are working on in order to achieve a more powerful and broad kind of AI. The results of this study should not be intended as unassailable evidence of the correlation between intelligence and creativity. Rather, they should be interpreted as a starting point for a discussion on how this interrelation unfolds and on how its analysis may promote the progress of research in both sectors.

While the results of the survey corroborate the work that is currently being done by AI researchers, in section 4 I put forward a second, tentative, hypothesis (H2): while the development of computational creativity may yield substantial advancements toward AGI, the opposite may not be true. Namely, the acquisition of features that are necessary to AGI may not be sufficient to design artificial systems that are recognized as creative.

A result that emerged from the survey is a generalized discontent in respect to the application of AI to the creative sector. If we assume H2 is true, then, we cannot explain this discontent through the current lack of general intelligence in state-of-the-art AI. I propose that the motivation that lies behind the audience’s disappointment could be explained through the uncanny valley theory. Although tentative, this argument can contribute to have a better understanding of the relationship between humans and technology when the latter enters a paradigmatically human field such as creativity.

## 2. ARTIFICIAL GENERAL INTELLIGENCE

Often the terms "Artificial General Intelligence" and "Human-level Artificial Intelligence" are used interchangeably as synonyms. However, this does not mean that their meanings match entirely. Indeed, while they both refer to a non-natural kind of intelligence, they direct the attention to different aspects of it. As Goertzel points out, the term "Human-level AI" is somewhat ill-defined, since it is hard to find a reliable tool to identify a hierarchy of natural and artificial kinds of intelligence. Loosely speaking, Human-level AI refers to an AI whose performance can match the standard performance of human intelligence. The concept of AGI, for its part, is more general and more fundamental. The notion of AGI is not, in fact, coupled with that of human-being, and it is not assessed in comparison to human performance.

AGI can be defined as a non-natural kind of intelligence "that can solve a variety of complex problems in a variety of different domains, and that controls itself autonomously, with its own thoughts, worries, feelings, strengths, weaknesses and predispositions." (Pennachin, Goertzel, 2007, p. 1) An AGI should learn to adapt to various environments and solve different problems in given situations, to appreciate the context in which it is situated, and to form connections between apparently different areas of application.

The term "AGI" gained widespread use in the field of AI after 2002, when Cassio Pennachin and Ben Goertzel used it as a title for their edited book which grouped contributions on approaches to AI systems that could solve broader and more complex problems than narrow kinds of AI (Pennachin, Goertzel, 2007, p. 2). The interest of researchers for the construction of a powerful intelligence with broad capabilities can be dated back to earlier times, though.

In the Fifties Allen Newell and Herbert Simon conducted several experiments toward the creation of general intelligence programs. It is the case of the General Problem Solver, a program designed in 1959 as a method to solve a variety of different problems, of the Logic Theorist, a program "capable of discovering proofs for theorems in elementary symbolic logic" (Newell, Shaw, Simon, 1962, p. 67), and of BACON, a program which re-discovered Kepler's Third Law and other laws of physics (Simon, 1985, pp. 9-10). Simon and his colleagues were confident of the bright future of AI and they shared the thought that, in a not so distant future, computers "should be able to take advantage of these capabilities to overtake humans: it was only a matter of a few years before suitable programming would let them do it." (Crevier, 1993, p. 5) Their predictions, however, turned out to be overly optimistic, so much so that 70 years later we still cannot say we have been able to build an AI that can reach human intelligence.

The aim of expanding the fields of application of AI and its potential is nevertheless very much alive and thriving in the AI community. This can be seen in the development of programs that can teach themselves how to play a variety of different games, like AlphaZero (Silver et al., 2017), in the latest innovations in the field of Natural Language Processing, with the creation of language models that can possess a deeper sense of the

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

context, like BERT (Devlin et al., 2019), and in the design of mobile robots that can navigate unstructured environments (Fabisch et al., 2019).

Still, it may be argued that, although impressive, these developments go all in the direction of a narrow kind of intelligence. AlphaZero, for example, cannot understand the nuances of a dialogue in a novel and Boston Dynamics robots cannot, for now, teach themselves how to play chess. The state-of-the-art text-generation program released by OpenAI in 2020, GPT3, is considered by many a radical step toward AGI.[1] And yet, although the program engages in an extremely complex task, such as that of generating human language, it still has not achieved the breadth and flexibility that characterizes human intelligence and behavior (Brown et al., 2020). Reactions to the AGI program range from the optimism of researchers like Goertzel all the way to the skepticism expressed by people like Yan LeCun or Andrew Ng and to a more worried attitude on the basis of the potential threat that it represents (Bostrom, 2014).

1 See https://www.theverge.com/21346343/gpt-3-explainer-openai-examples-errors-agi-potential.

I will not enter the debate on the hopes and concerns posed by the progress of AI, since this would go beyond the aims of this paper. Still, while the development of more powerful and broader capabilities of AI undeniably raises ethical and societal questions which are worth being discussed (Yampolskiy, Fox, 2013), we are still far from the design of system that can reach, let alone surpass, human intellectual abilities.

In section 3, I will discuss how the challenges researchers meet in programming AGI systems are correlated to aspects that are essential also to develop artificial creativity. To support this claim, I draw on the results obtained from a survey on the public reception of creative artificial systems that I present in the next section.

## 3. SURVEY ON COMPUTATIONAL CREATIVITY

Since the attention brought to the field by Guilford's speech on creativity (Guilford, 1950), many studies have focused on the nature of creativity and on its application to various domains.[2] The question of whether, alongside humans and other animals, also artificial systems can be creative has gained more attention since the Nineties, with the work on computational creativity by Margaret Boden (2004). In the last decades, many are the programs designed with the aim to build systems that exhibit creativity in visual arts (Colton, 2012), music (Eigenfeldt, Pasquier, 2011), poetry (Gatti et al., 2012), and even fashion and cooking (Morris et al., 2012).

2 For an overview on the theme of creativity in the literature and on the different definitions that can be given of it, see Boden, 2004; Elton, 1995; Gaut, 2010; Glover, Ronning, Reynolds, 1989; Moruzzi, 2020b; Newell, Shaw, Simon, 1962; Runco, Jaeger, 2012; Keith Sawyer, 2012; Simon, 1985; Sternberg, 1999; Weisberg, 1993.

The scope of this analysis is not to assess the technical performance achieved by state-of-the-art technologies in the creative sector. Rather, the aim is to examine the opinions and perceptions that scholars, creatives, and the general public have towards AI that produces "Art". To this end, in this section I discuss the results obtained from a study conducted at the University of Nottingham.

The study "Perceptions of Creativity in Human and Artificial Intelligence" has been conducted in September 2019 as part of the research priority area project "Audience Perceptions of AI Interaction in

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

Different Modes of Engagement", in collaboration with the Department of Humanities and Medicine of the University of Nottingham. [3] The survey was intended as a limited, preliminary study designed to raise questions and pave the way for further research on the topic. It included two parts: an online questionnaire, completed by 203 participants, and two focus groups, with 10 participants in each group and a duration of one hour each.

A large group of participants of both the online questionnaire (67.5%) and the focus group (25%) came from an academic background or were still undertaking their studies (33% for the online questionnaire, and 35% for the focus group). The limitations of the study derive then quite clearly from the small and unrepresentative sample group. To have a more statistically relevant sample it would, therefore, be recommended to carry out a subsequent study that focuses on a different sector of the population.

After a screening section, aimed at measuring the familiarity participants had with AI and their engagement with the art sector, participants were asked their opinion regarding the application of AI in a range of different sectors. The central part of the questionnaire asked participants to indicate the principal features of creativity and to give answers to targeted questions about the use of AI in the creation of supposedly artistic products. Lastly, the questionnaire closed with some general questions on the participants' opinion regarding the possibility for AI to be creative.

What emerges from the replies given by participants in the questionnaire is a generalized disappointment in respect to the creation of artistic products by AI systems. To the questions "If you found out that the painting/music album you just bought was not painted/composed by a human but by an AI, how would you feel about it?", participants did not react with much enthusiasm. The 31% of participants would be positively surprised if the painting was created by an AI and 35.5% if the music album was composed by an AI, but almost the same number of participants would react in a neutral way (36.5% for the music album and 36.9% for the painting). The 28.1% and 32% would instead be disappointed in finding out that the music album and painting respectively were AI-generated. Even clearer is the participants' opinion when asked "If you had the choice of buying a painting created by a human and one created by AI, which one would you buy?". The 90.1% of participants would buy a painting created by a human and the 93.1%, when asked the same question but in relation to a music album, would prefer to buy an album composed by a human.

The central task that the questionnaire presented to participants was to answer some questions about two paintings and two musical clips. Participants were not aware of the fact that one of the two paintings was created by an AI and one of the two clips was AI-generated. Despite the disappointment expressed by participants in respect to the AI-generated painting and music given as example in the test, [4] the majority of participants agreed with the possibility for AI to be creative: 40.9% of participants replied "Yes", and 11.8% replied "Certainly". Even more

3 The other members of the team that participated in the project are Dr. Elvira Perez Vallejos, School of Medicine, Dr. Nicholas Baragwanath, Department of Music, Dr. Zachary Hoskins, Department of Philosophy.

4 Unaware of their provenance, 40.9% of participants agreed that Painting 1 (human-generated) was creative or very creative, while 28.6% claimed that Painting 2 (AI-generated) was creative or very creative. As for the clips, 35% of participants agreed that Clip 2 (human-generated) was creative or very creative, while only 20.2% claimed that Clip 1 (AI-generated) was creative or very creative. The results obtained are similar also in respect to other parameters: the human-generated painting and clip obtained higher percentage of confidence as for their novelty, pleasantness, and surprisingness and participants liked them more than the AI-generated painting and clip.

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

optimistic is the reaction to the question “Do you think humans can be more creative with the help of AI?”. In this case the 44.3% replied “Yes” and the 29.6% is certain of the benefit that may come from the collaboration between humans and AI.

The details of the survey and of its results are reported elsewhere (Moruzzi, 2020a). In this paper, I focus on what emerged from the results of the survey in relation to the following hypothesis: H1: in order to acquire creative abilities, AI systems need to develop features that are necessary also to progress from Narrow to General AI. In what follows I am going to discuss how the features that participants reported as lacking in state-of-the-art artificial systems and that, instead, would be necessary for them to be deemed creative, correlate with some of the same crucial aspects that currently represent a challenge to researchers in the development of AGI.

## 4. CREATIVITY AND GENERAL INTELLIGENCE

It seems hard to deny that creativity and intelligence are deeply interrelated. What is more contentious, however, is what exactly the correlation between the two concepts amounts to. This difficulty is made even harder by the fact that both are polysemous terms which are hard to demarcate and define.

The notion of creativity can indeed cover a wide spectrum of meanings and definitions. It can be described as a subjective property of the artist or as a quality that is assigned to the process or product in question by the audience. On the other hand, it can also be described as an objective property that can be developed through exercise and hard work (Simon, 1985).

Intelligence, for its part, can be defined, for example, as the capacity to acquire and apply knowledge, as a more general ability to solve problems, or as a more fundamental capacity to adapt to the context (Legg, Hutter, 2007).

If one digs deeper into the variety of ways in which the two concepts have been described by researchers, it is possible to find many parallels and overlaps. The ability to solve novel problems, for example, is a capacity which is recognized by many as both a mark of intelligence and as a key aspect of creativity (Simon, 1985; Sawyer, 2012; Weisberg, 1993). Others consider computational creativity as a tool to endow artificial systems with the necessary robustness to progress towards a more general kind of intelligence (Ventura, 2014, p. 1). The aim of the following discussion is precisely that of exploring how the development of creative features can be beneficial towards a progress in intelligence and viceversa, focusing on the aspects that have been highlighted by participants on the survey.

The concerns in relation to the use of AI systems for the creation of artistic products is expressed by many participants with clarity - and also with some vehemence - in reply to the open question of the survey. The question asked participants to motivate their reply in response to the

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

question “How likely do you think it is that AI replaces human artists in the next 10 years?”. 176 out of the 203 participants replied extensively to the open question, despite it being not compulsory for moving on with the questionnaire, thus showing the wish to engage with the topic.

The majority of replies express skepticism about the possibility for AI to become creative: AI systems cannot generate creative products/be creative, since they lack experience and contextual knowledge, flexibility, feelings, intentionality, and other features, such as unpredictability and the ability of self-evaluation.

Some of the participants' replies identified domain knowledge and the ability to transcend it as necessary elements of creativity. On the one hand, in fact, knowledge and experience in the field and the capacity to successfully engage with the environment are necessary to a creative system and it is this breadth of experience that AI is deemed to presently lack:

> 4130 — “Interesting art usually presupposes a lot of knowledge of the history of art.”
> 5617 — “Good art requires ability and skill and a working knowledge of the artistic tradition but also some kind of intention in the sense of insight or a message or a theme that is being communicated. It is about something. In all probability, AI will produce kitch that lacks substance.”
> 0888 — “Intention is besides the point (so too ‘pleasant’ and ‘surprising’) - the age itself speaks through the work of art. How is AI to acquire history? Context?”
> 0177 — “Intuition is a refined form of emotional intelligence which whilst it seems ‘easy’ is actually highly complex and dependant (sic) entirely upon a breadth of experience which AI will not be able to replicate.”

The acquisition of domain-knowledge is an aspect which is deemed fundamental also by researchers in AGI. One of the methods used in order to achieve this aim is to encode in artificial systems a knowledge-base containing all human common-sense knowledge (Pennachin, Goertzel, 2007, p. 3).

Yet, the possession of broad knowledge in the field at hand is not sufficient to develop general intelligence. Successfully operating in a common sense informatic situation is a requisite for developing AGI as well (McCarthy, 2007, pp. 1175-76). A common sense informatic situation is opposed to a bounded informatic situation. It is more general than the latter and it requires flexibility and adaptation in understanding and engaging with the environment (Goertzel, 2004, p. 41). This flexibility is related to the “insight” and “intuition” participants talk about in relation to creativity and it is, if possible, even harder to instill into artificial systems than broad domain-knowledge.

According to many, in order to be creative, a system does not only need to possess domain-knowledge but also the ability to naively transcend it (Grba, 2020; Ventura, 2014, p. 3). The same is arguably true

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

for intelligence (McCarthy, 2007, p. 1178; Pennachin, Goertzel, 2007, pp. 6-7): "Intelligence is the ability to work and adapt to the environment with insufficient knowledge and resources." (Pennachin, Goertzel, 2007, p. 10)

The notion of naivety interrelates with aspects such as spontaneity (Dewey, 1934; Kronfeldner, 2009; Sawyer, 2012), unconscious thought processing (Baumeister et al., 2014), and independence from rigid structures of thought, all of which have been listed in the literature among the core traits of creativity. Describing the notion of naivety as independence from a model and from external causal inferences does not have the consequence of denying the relevance of domain awareness and expert knowledge (Glover, 1989; Jordanous, Keller, 2016; Sawyer, 2012; Simon, 1985; Weisberg, 1993), rather it recognizes that a creative process is a process of exploration which does not necessarily demand expertise and self-education.

A further aspect that participants considered necessary to creativity is the capacity of the agent of self-evaluating its process and products. This capacity of introspection, however, is deemed as presently lacking to AI systems:

> 2916 — "AI have no way of judging if what they have created necessarily sounds/looks good."

The ability to assess the process and to "know when to stop" is a relevant element of creativity (Moruzzi, 2020bb; Boden, 2004, pp. 43-44; Gaut, 2010; Guckelsberger et al., 2017; Rhodes, 1961, p. 305). The evaluative component of creativity is crucial in the process of trial and error that is carried out by agents when engaging in creative endeavors (Sawyer, 2012, pp. 107-110; Weisberg, 1993, p. 109). This process of self-assessment should not be a result of external influences of feedback, but rather an autonomous process.

Self-awareness and the autonomy of the system from external feedback are essential not only to develop creativity but also to achieve a human-level intelligence (Goertzel, 2004, p. 9; McCarthy, 2007, p. 1178; Ventura, 2014, p. 6). Introspection is necessary for an agent to learn and to adapt to the environment. Indeed, intelligent agents adjust their behavior to generate an output that is as much adherent as possible to their initial goal.

A last element I consider in this analysis is the relevance of embodiment for both creativity and intelligence. In the second part of the focus group, I showed participants the music video "Automatica" of the musician Nigel Stanford playing music together with robotic arms created by the company KUKA Robotics.[5] The video had the aim of introducing a further variable in the discussion, namely a physical presence of the machine that actively performs on stage.

Despite the embodied presence of the robots - an element that would supposedly contribute towards the anthropomorphization of the machines (Goldman, 1993; Edmonds, 1994; Sharples, 1994) - participants did not perceive them as real "musicians" or "players". Rather, they

5 Available at https://www.youtube.com/watch?v=bAdqazixuRY.

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

reported a lack of engagement with the performance and "with something so different from us" (Participant 6). The lack of interaction with the audience led a participant to assert that he/she would "rather listen to the music but not see the robots" (Participant 9).

Even if the robotic arms had not been deemed sufficient to vouch in favor of a display of creativity of the system, though, the relevance of embodiment for artistic performances had been acknowledged. Indeed, some participants did not exclude the possibility for these artificial agents to build a better connection with the public if they develop and assume additional "human-like" physical features. Embodiment is, indeed, a relevant aspect of both creativity and intelligence:

> 5487 — "Human artists have a face and a body, they can play a real instrument, they can be talented, and can attract fans for all these aspects. No AI will have this sort of carisma [sic]"

Even if state-of-the-art AI may not have been able to impress participants, embodiment approaches are undertaken by researchers to advance toward AGI (Goertzel, 2004, p. 13). Intelligence develops in physical bodies and in physical environments, as agents need to adapt to environmental constraints and to interact with the environment. The possibility for artificial systems to possess an embodied presence is key to progress toward a more general and flexible kind of intelligence.

In part, the opinions expressed by participants and discussed above confirm the hypothesis that, in order to be creative, AI systems need to develop aspects that researchers are currently working on in order to progress towards building AGI. Intuition, unpredictability, and self-awareness, features that participants indicated as currently lacking in AI systems, are elements necessary not only for an agent to be creative, but also for AI to reach a broader and more flexible intelligence[6].

In this respect, the survey and its results corroborate the research conducted in the field of AGI. However, other elements that were not assumed by H1 and that, instead, are deemed essential for the development of creativity by participants, emerged. In the next section I discuss which other aspects an artificial agent would seem to require in order to be deemed creative, discussing at the same time one of the possible reasons that may stand behind the participants' discontent against AI.

## 5. ANTHROPIC AI AND UNCANNY VALLEY

Alongside the features above mentioned, what participants reported as lacking in state-of-the-art AI, and that would instead be necessary for them to be deemed creative, are aspects that are essentially human, such as charisma, personality, experience, and the property of having and transmitting feelings and emotions. What emerges from the participants'

6 It may be argued that a survey of non-experts is not the right kind of evidence to prove or disprove this hypothesis. Still, these results may open to further research paths, and can prove to be relevant when designing creative AI, as I also indicate in section 4.

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

replies is the belief that creativity and art are human-centric domains that are, and should remain, the prerogative of humans:

> 8215 — “The human subjectivity, intentionality and creativity can’t be replaced by a machine, because a machine can only imitate the objective-formal thinking process.”
> 3148 — “Pieces of art do not just represent something pretty or novel, but are also a representation of our lives and our reality as humanity throughout the years. […] An AI would probably never be able to have such empathy and views of the world as an artist to represent through it’s [sic] art those issues.”
> 3776 — “Art will remain a human domain because AI is not able to have qualitative experience […] only if you taste sadness, you can know how to express it in a way that can be sharable with other human beings.”
> 3049 — “Art is an expression of the artist’s innermost thoughts, feelings and beliefs and I highly doubt AI will ever be able to achieve this.”
> 2916 — “Coming up with new ideas and expressing creativity is something hard to replicate, it’s uniquely human.”

I, thus, put forward a second, more tentative, hypothesis: H2: the achievement of an artificial intelligence that possesses the features that characterize general intelligence would not suffice to make people accept the possibility for AI to be creative.

The reports mentioned above seem to provide evidence in favor of H2: the acquisitions of capacities that characterize general intelligence, such as common-sense reasoning, flexibility, self-evaluation do not suffice to ascribe to AI systems the attribute of “creative”. This assumption is in line with the argument according to which, while creativity is a necessary element of intelligence, intelligence on its own is not sufficient for creativity.[7] Something else is needed and the participants’ comments seem to lead us toward the conclusion that these additional requirements are anthropogenic ingredients.

Assuming H2 is correct, and not even the acquisition of features that are key to general intelligence would suffice to make people recognize AI as creative, the lack of these features in state-of-the-art AI does not provide a sufficient explanation for the participants’ disappointment against creative AI. In what follows, I tentatively propose to find a motivation for this discontent by interpreting it within the framework of the uncanny valley and through the human process of unconscious creation of mental models.

The uncanny valley is a theory proposed by the robotics professor Masahiro Mori in 1970, which argues for a non-linear relation between the degree of human-like appearance and likeability of a robot. In this theory, Mori describes how “in climbing toward the goal of making robots appear like a human, our affinity for them increases until we come to a valley”, what is called, indeed, the *uncanny valley* (Mori, 2012, p. 98). As

7 See https://creativesomething.net/post/41103661291/the-relationship-between-creativity-and.

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

soon as we realize that, what looked real, is artificial, we experience an uncomfortable feeling of eeriness.

Mori tries to explain the reason behind the emergence of the human reactions associated with the perception of entities which fall into this valley through the innate human instinct of self-preservation which protects humans from proximal sources of danger (Mori, 2012, p. 100).

The phenomenon of the uncanny valley is rooted in the more general concept of anthropomorphism. It is widely acknowledged that humans have an ingrained propensity for anthropomorphizing entities and objects of their environment, from rocks, to clouds, to robots (Guthrie, 1993). Epley and his colleagues explain this tendency with the need to better understand and control the environment and to create social bonds (Epley et al, 2008). Within the context of the human-machine relation, we can find an example of this latter motivation in the connection that humans look for in the interaction with chatbots, care robots, or even sex robots (Sparrow, Sparrow, 2006). On the other hand, the difficulty that we may experience in understanding and interpreting state-of-the-art AI systems is at the source of many of the fears towards apocalyptic scenarios where AI takes over humans (Bostrom, 2014).

Anthropomorphization does not necessarily occur only in relation to physical features and humans may anthropomorphize other entities also on the basis of unobservable characteristics (Epley et al., 2008, p. 144; Gray and Wegner, 2012). In the same way, the phenomenon of the uncanny valley does not happen only in relation to human-appearing robots but also to artificial entities with a more abstract appearance (Walters et al., 2008, p. 160). Moreover, the unnerving feeling associated to the uncanny valley may occur not only in relation to the degree of human-like appearance of a machine. Rather, also the perception of experiences in a machine may underlie this phenomenon. The possibility of having experiences is judged as an essentially human characteristic: “we are happy to have robots that do things, but not feel things” (Gray, Wegner, 2012, p. 129; Airenti, 2015, p. 125) What is unsettling about a machine that appears to be capable of love, fear, hate, excitement, etc. is that it violates expectancies of experience. Machines apparently capable of experience are unnerving, even without a human-like appearance (Gray, Wegner, 2012), because we do not accept to attribute this capacity to them.

People tend to form a mental model of an entity by making assumptions from its features. This mental model serves as a guide to the performance that should be expected from that entity. Two scenarios may occur in this respect: (i) if the appearance is more advanced than the entity’s performance, then humans tend to judge that entity as “dishonest”; (ii) if, instead, the appearance is less advanced than the performance, people do not fully exploit the entity’s capabilities (Walters, 2008, p. 161).

As discussed, creativity is a human-laden concept which requires features such as personality and intentionality. As can be derived from the reports, people possess a mental model of machines and artificial entities in general which does not include the possibility for them to experience,

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

Journal of Science and Technology of the Arts, Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

nor to possess other, paradigmatically human, features such as creativity (Gray and Wegner, 2012). The fact that, nevertheless, artificial entities instantiate behaviors and activities that are typical of creative agents, is perceived as dishonest or, even more drastically, it triggers uncomfortable feelings of rejection and discontent against the entities in question.

This is exemplified by the audience's response to the "Automatica" video described above. The robots in the music video instantiated movements and behaviors typical of musicians. The mental model that the audience had of robots, namely artificial entities without neither experience nor charisma, clashed with the involvement in a creative activity that the robots seemed to embody on stage. This clash triggered the disappointment of the participants who felt somehow "cheated", as they failed to attribute human features such as personality to the robots, although the latter were acting "as if" they possessed them.

The discontent expressed by participants to the survey against AI engaging in creative activities can, therefore, be tentatively explained through a process similar to the one occurring in the uncanny valley phenomenon. It is a widespread belief that creativity is characterized by essentially human features and, in parallel, the mental model that we have of artificial entities does not include the possibility for it to be creative. When artificial entities engage in creative activities and stage stereotypically human features which they are not deemed capable to really possess, they fall into a valley of disapproval and disappointment which distances them further from us, instead of increasing the affinity that we may have with them. The fact that AI takes part to creative activities is disturbing and unnerving because it unconsciously would lead us to ascribe to it creativity, a property that people arguably are not willing to attribute to artificial entities (Airenti, 2015, p. 124).

The former should not be intended as a definite answer to what the motivations for the attitudes people have toward creative AI are. However, it might contribute to the debates on human-machine interaction by raising questions and proposing tentative answers to the mechanisms that lie behind this relationship. Moreover, by interpreting the public discontent against creative AI in the light of the uncanny valley phenomenon, this work contributes to the discussions regarding the validity of Mori's theory in different sectors of AI research.

Mori's warning, I argue, is still valid today and it needs to be considered when designing artificial systems that engage in creative activities (Salles et al. 2020). This is crucial to prevent the risk of not taking advantage of the benefits that may derive from the human-machine interaction in the creative sector due to biases or pre-determined mental models that we humans have formed.

## 6. CONCLUSIONS

In the Seventies, Mori warned against the danger of the uncanny valley in our quest toward the achievement of a Human-level intelligence. I believe that the results of the study presented in this paper, although preliminary, corroborate the validity of this warning. While acknowledging the relevance of the work researchers are doing toward the development of general intelligence in artificial systems, the participants' replies point out how this is not enough in order to attribute creativity to them. The achievement of features such as flexibility, naivety, embodiment, and evaluation – essential for general intelligence – arguably would not suffice to make humans accept artificial entities that engage with paradigmatically human activities, such as creativity. What motivates the audience's disapproval is not the fact that state-of-the-art AI still exhibits a 'narrow' kind of intelligence. Rather, this irritation is motivated by the fact that artificial entities pretend to possess something that they do not have (yet): the capacity of feeling and experiencing like a human. Thus, creative AI falls into a sort of uncanny valley, distancing even more from our approval and sympathy.

Still, while the development of AGI may not necessarily lead to the development of creative AI, the opposite is less contentiously true: research in the field of computational creativity may yield substantial advancements toward AGI (Ventura 2014). The investigation on the nature of creativity and on how it manifests itself not only in human but also in animal and artificial systems should, thus, not be intended as a niche discussion but, rather, as a fundamental research which can lay the foundations for further studies in artificial intelligence and its relation to humans.[8]

8 I thank two anonymous reviewers for their comments on a previous version of this manuscript.

## 7. REFERENCES

Airenti, G. (2015). The Cognitive Bases of Anthropomorphism: From Relatedness to Empathy. *International Journal of Social Robotics 7*, 117–127.

Baumeister, R.F. et al. (2014). Creativity and Consciousness: Evidence from Psychology Experiments. *The Philosophy of Creativity: New Essays*, eds. E.S. Paul and S.B. Kaufman. Oxford University Press.

Boden, M. (2004). *The Creative Mind: Myths and Mechanisms* (2nd ed.). George Weidenfeld and Nicolson.

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

Brown, T. et al. (2020). *Language Models are Few-Shot Learners*. arXiv:2005.14165v4.

https://doi.org/10.34632/jsta.2020.9481

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99

Colton, S. (2012). The Painting Fool: Stories From Building an Automated Painter. In J. McCormack & M. D'Inverno (Eds.), *Computers and Creativity* (pp. 3-38). Springer.

Crevier, D. (1993). *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books.

Devlin, J. et al. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv:1810.04805v2.

Dewey, J. (1934). *Art as Experience*. New York: Minton, Balch & Co.

Edmonds, E. (1994). Introduction: Computer-based Systems that Support Creativity. In T. Dartnall (ed.), *Artificial Intelligence and Creativity, Studies in Cognitive Systems* (pp. 327-334).

Eigenfeldt, A. & Pasquier, P. (2011). Negotiated Content: Generative Soundscape Composition by Autonomous Musical Agents in Coming Together: Freesound. *Proceedings of the 2nd International Conference on Computational Creativity*, 27-32.

Elton, M. (1995). Artificial Creativity: Enculturing Computers. *Leonardo 28*, 207–213.
*https://doi.org/10.2307/1576076*

Epley, N. et al. 2008. When We Need a Human: Motivational Determinants of Anthropomorphism. *Social Cognition 26*(2), 143-155.

Fabisch, A. et al. (2019). *Boston Dynamics, A Survey of Behavior Learning Applications in Robotics - State of the Art and Perspectives*. arXiv:1906.01868v1.

Gatti, L., Guerini, M., Callaway, C., Stock, O., Strapparava, C. (2012). Creatively Subverting Messages in Posters. *Proceedings of the 3rd International Conference on Computational Creativity*, 175-179.

Gaut, B. (2010). The Philosophy of Creativity. *Philosophy Compass 5*, 1034-1046.
*https://doi.org/10.1111/j.1747-9991.2010.00351.x*

Glover, J., Ronning, R., Reynolds, C. (Eds.). (1989). *Handbook of Creativity*. Plenum Press.

Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence 5*(1), 1-46.
*https://doi.org/10.2478/jagi-2014-0001*

Goldman, A. I. (Ed.). (1993). *Readings in Philosophy and Cognitive Science*. MIT Press.

Gray, K., Wegner, D. M. (2012) Feeling Robots and Human Zombies: Mind Perception and the Uncanny Valley. *Cognition 125*, 125-130.
*http://doi.org/10.1016/j.cognition.2012.06.007*

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

Grba, D. (2020) Deriving Sense: Cognitive Aspects of Artefactual Creativity. *Proceedings of xCoAx2020*, 186-201.

Guckelsberger, C. et al. (2017). Addressing the “Why?” in Computational Creativity: A Non-Anthropocentric, Minimal Model of Intentional Creative Agency. *Proceedings 8th International Conference of Computational Creativity*.

Guilford, J.P. (1950). Creativity. *American Psychologist 5*, 444-454. *https://doi.org/10.1037/h0063487*

Guthrie, S. E. (1993). *Faces in the Clouds: A New Theory of Religion*. Oxford University Press.

Jordanous, A. & Keller, B. (2016). Modelling Creativity: Identifying Key Components through a Corpus-Based Approach. *PLoS ONE*. *https://doi.org/10.1371/journal.pone.0162959*

Kronfeldner, M. E. (2009). Creativity Naturalized. *The Philosophical Quarterly 59*, 577-592. *https://doi.org/10.1111/j.1467-9213.2009.637.x*

Legg, S., Hutter, M. (2007). *A Collection of Definitions of Intelligence*. arXiv:0706.3639.

McCarthy, J. (2007). From Here to Human-level AI. *Artificial Intelligence 171*, 1174-1182. *https://doi.org/10.1016/j.artint.2007.10.009*

Mori, M. (2012). The Uncanny Valley. *IEEE Robotics & Automation Magazine* *https://doi.org/10.1109/MRA.2012.2192811*

Morris, R. et al. (2012). Soup Over Bean of Pure Joy: Culinary Ruminations of an Artificial Chef. *Proceedings of the 3rd International Conference on Computational Creativity*, 119-125.

Moruzzi, C. (2020a). Should Human Artists Fear AI? A Report on the Perception of Creative AI. *Proceedings of xCoAx2020*, 170-185.

Moruzzi, C. (2020b) Measuring Creativity: An Account of Natural and Artificial Creativity. *European Journal for Philosophy of Science*, *11*(1).

Newell, A., Shaw, J.C., Simon, H. (1962). The Process of Creative Thinking. In H. Gruber, G. Terrel, M. Wertheimer (Eds.), *Contemporary Approaches to Creative Thinking* (pp. 63-119). Atherton Press.

Pennachin, C., Goertzel, B. (2007) Contemporary Approaches to Artificial General Intelligence. In: Goertzel B. & Pennachin C. (Eds.), *Artificial General Intelligence. Cognitive Technologies*. Springer.

Rhodes, M. (1961). An Analysis of Creativity. *Phi Delta Kappa 42*(7), 305-310.

**Journal of Science and Technology of the Arts,** Vol. 12, n. 3 (2020): pp. 84-99 https://doi.org/10.34632/jsta.2020.9481

Runco, M. A. & Jaeger, G. J. (2012). The Standard Definition of Creativity. *Creativity Research Journal 24*, 92-96.
*https://doi.org/10.1080/10400419.2012.650092*

Salles, A. et al. (2020). Anthropomorphism in AI. *AJOB Neuroscience 11* (2), 88-95.
*https://doi.org/10.1080/21507740.2020.1740350*

Sawyer, R. (2012). *Explaining Creativity: The Science of Human Innovation*. Oxford University Press.

Sharples, M. (1994). Cognitive Support and the Rhythm of Design. In T. Dartnall (Ed.), *Artificial Intelligence and Creativity: Studies in Cognitive Systems*, 385-402.

Silver, D. et al. (2017). *Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm*. arXiv:1712.01815.

Simon, H. (1985). What We Know About the Creative Process In Robert Kuhn (Ed.), *Frontiers in Creative and Innovative Management* (pp. 3-22). Ballinger.

Sparrow, R. & Sparrow, L. (2006). In the hands of machines? The future of aged care. *Minds and Machines, 16*(2), 141–161.
*https://doi.org/10.1007/s11023-006-9030-6*

Sternberg, R. J. ed. (1999). *Handbook of Creativity*. Cambridge University Press.

Ventura, D. (2014). Can a Computer be Lucky? And Other Ridiculous Questions Posed by Computational Creativity. In B. Goertzel et al. (Ed.), *Proceedings of 7th International Conference AGI 2014*, pp. 208–217.

Walters, M. L. et al. (2008). Avoiding the Uncanny Valley: Robot Appearance, Personality and Consistency of Behavior in an Attention-seeking Home Scenario for a Robot Companion, *Autonomous Robots 24*, 159–178.

Weisberg, R. (1993). *Creativity: Beyond the Myth of Genius*. W.H. Freeman.

Yampolskiy, R. & Fox, J. (2013). Safety Engineering for Artificial General Intelligence. *Topoi 32*, 217–226.
*https://doi.org/10.1007/s11245-012-9128-9*

Article received on 11/08/2020 and accepted on 11/11/2020.