# Every dog has its day: An in-depth analysis of the creative ability of visual generative AI

MARIA M. HEDBLOM

**Abstract:** The recent remarkable success of generative AI models to create text and images has already started altering our perspective of intelligence and the "uniqueness" of humanity in this world. Simultaneously, arguments on why AI will never exceed human intelligence are ever-present as seen in Landgrebe and Smith (2022). To address whether machines may rule the world after all, this paper zooms in on one of the aspects of intelligence Landgrebe and Smith (2022) neglected to consider: creativity. Using Rhodes four Ps of creativity as a starting point, this paper evaluates the creative ability in visual generative AI models with respect to the state of the art in creativity theory. The most part of the reflective evaluation is performed through a case study in generating illustrations of dogs using the generative AI tool Midjourney.

## 1. SOMEWHERE BETWEEN BLASPHEMY AND DIVINE INTERVENTION

In Cologne Cathedral one of the stained-glass windows stands out. Unlike traditional church windows with Biblical scenes, religious figures, floral patterns and other expressions of God's creation, the window lacks motif and instead consists of a squared grid of colored cells. Created by contemporary artist Gerhard Richter, the window's colorful pixel-like constellation is the result of implementing and prompting the "random function". Since its reveal in 2007, the art piece has received a mixture of both praise and criticism: argued to both be a meditative product due to its lack of message and blasphemously as the underlying process that created the product did not have any conscious intent (Belz 2007).

The Richter-window is an early example of the introduction of computer-generated visual art into the cultural scene of human society. Whether people find the window aesthetically pleasing or not, is not per se a reflection of its creative character, but few people would probably argue that a system randomly placing color pixels in a grid formation displays any level of conscious creative effort. Instinctively, humans require there to be some thought, or at the very least some serendipitous happenstance, behind the generative process of an artistic product for it to be considered creative. Following this line of reasoning, the reactions to the window display an interesting phenomenon. Namely, the reluctance humans have to ascribe creative ability to a non-human entity. Within the setting of artificial intelligence (AI), Colton and Wiggins compare creativity to the perception of intelligence by saying that "...perhaps creativity is, for some proponents of AI, the place that one cannot go, as intelligence is for AI's opponents. After all, creativity is one of the things that makes us human; we value it greatly, and we guard it jealously" (Colton and Wiggins 2012, p.21).

This reasoning runs parallel to arguments on the requirements of a system for it to be deemed intelligent. In Searle (1980)'s famous thought experiment The Chinese Room, the argument goes that if a system is simply transforming symbolic input into symbolic output by following a series of transformation rules, the system is claimed to not understand the task nor to have any semantic grounding of the involved symbols. The main idea behind the argument is that the system lacks intentionality[1] and the ability to ground the semantics of the transformation. Essentially, the Richter-window follows a similar process: colors and grid as input, a random function that assign color to the cells in the grid, and a constellation of colors in the grid as output.

Today, the computational techniques behind the Richter-Window are rather outdated as the technological advancement of statistical methods such as deep learning (DL) has revolutionized research in AI to the point in which systems have started to display what on surface level appears to be human-level competence in certain areas. Using big data and transformer models, large language models (LLMs) (e.g. Google's BERT (Devlin et al. 2018) and openAI's GPT3)[2] have demonstrated remarkable skills in replicating human language and can provide contextually appropriate answers to questions in almost any topic.

This remarkable success has had many researchers, practitioners and laypeople speaking of AI as on their way to reaching 'artificial general intelligence' and that synthetic phenomenology[3] should be treated as a form of machine consciousness (Metzinger 2021). In their book, (Landgrebe and Smith 2022) call for a more critical view of the level of general intelligence in modern AI approaches by arguing that artificial general intelligence is a 'mathematical impossibility'. Their argument is that mathematical modeling, used in (most) modern AI approaches, is fundamentally unable to capture the complexity and variety of animated behavior. Not claiming to fully comprehend the extent of Landgrebe and Smith's (2022) argumentation and yet agreeing with the counterarguments by Rapaport (2023) on the limits of their base assumptions, it is reasonable to take a critical account of the current state of AI capacities. One topic that was not discussed in the work by (Landgrebe and Smith 2022) was what Colton and Wiggins (2012) argued to be one of the central components of human intelligence, namely creativity. Thus, in order to contribute to the ongoing AI debate, this contribution takes a closer look at how modern AI fares in relation to creativity theory, in particular, in relation to the nature of artistic expression.

For modern AI approaches dealing with artistic creativity, the same methodological principles underlying LLMs can be seen in image-generative AI. These systems are trained on text-to-image pairs that, just like their text-based counterparts, can generate remarkable images and pieces of art from text prompts (e.g. systems like DALL-E,[4] Stable Diffusion[5] and Midjourney)[6]. Like with LLMs, these systems are predictive and generative models that are trained on enormous amounts of data to 'fill in the blanks' of missing pieces of images and generate visual responses to prompted requests.

*(a) Stable Diffusion* *(b) DALL-E2* *(c) Midjourney*

*Figure 1: Generated images with the prompt "praying hands." The images were selected from the first generated batch and not modified in any way.*

One important thing to note about these generative AI models is that the architecture of these computational artists does not have access to any method for explicit semantic grounding. This means that when they are prompted with a word like 'dog', the system has no internal representation of a dog nor access to some conceptual categorization of the concept. Instead, trained on images of millions of dogs, the system recreates a 'mean' of these images by choosing the color of pixels based on principles of the closest neighbors. While also humans have a sort of visual 'prototypical dog' in our minds (Hampton, 1993)—that may or may not look like any dog in particular, we also attach this visual representation with a conceptual space of all that we know about 'dogness'. For generative AI, there is no commonsense knowledge or conceptual space involved in the generation. This means that the generation can result in rather absurd and unrealistic representations. One commonly mentioned example is the multitude of 'wrongs' that can materialize when these systems are asked to generate hands. To demonstrate this, Figure 1 shows three generated pictures with the prompt "praying hands". While extraordinary images, take note of the placement and number of fingers. With problems such as this in mind, humans retained their feeling of superiority in artistic generation. Then in 2022, the unthinkable happened. "Théâtre D'opéra Spatial", an image generated by Midjourney prompted by Jason Allen, won first prize in the digital category at the Colorado State Fair's art competition. This means, with or without jealously (Colton and Wiggins 2012), it is no longer possible to speak of AI-generated art as any less aesthetically pleasing than that created through human expression. Further, in relation to the ongoing debate in this paper, we have to question Landgrebe and Smith's (2022) argument that "mathematics is not enough" for artistic creativity.

Returning to the Chinese Room argument, the influential paper by Bender and Koller (2020) provides an interesting angle to the intentionality problem of generative AI. They argue that training LLMs on large amounts of data teaches the model *form* not *understanding*.[7] While the disjoint distinction between form and meaning is an important issue to discuss and (ideally) solve for AI dealing with text generation, a liberal interpretation of artistic creativity could argue that for AI generating art—where *form* is the goal—any system that successfully generates form should be deemed to demonstrate the creative ability to generate art.

Creativity is in itself not easier to define than art, but it does offer the possibility of a more systematic analysis and comparison between creativity expressed in humans and that of computer systems. In the remainder of this article, AI-generated creativity will be analyzed based on perspectives and metrics from creativity theory.

## 2. ARTISTIC CREATIVITY: FROM PALEOLITHIC STATUES TO GENERATIVE AI

Humans are reluctant to ascribe creative ability to any non-human entity. This is not exclusively seen in the area of AI and computer-generated imagery but a parallel can be made to the view on creative ability in animals. Such expressions have been argued to predominantly express evolutionary beneficial behavior. Thus, artistic and creative problem-solving have been reserved for humans. Within this 'evolutionary' setting, a male pufferfish calls to females by creating beautiful, mandala-like patterns in the sand, male birds of paradise perform ritualistic mating dances and animals such as great apes display a relatively high level of innovative problem-solving capabilities (Manrique et al. 2013). Artistic expression and creative problem-solving might be particularly prominent and exceptional (based on our own evaluation criteria) in human expression, but humanity is by no means the sole possessor of creative abilities.

What distinguishes humanity is our (superior) ability for mental abstraction and using art as symbolic representations. Throughout human evolution, the development of complex language brought with it an increased sophistication in creative ability by being able to enrich creative expression with symbolic meaning and attribution (Morriss-Kay 2010). This can, for instance, be seen in the oldest known man-made sculpture The Lion Man, estimated to be some 35000-41000 years old. In one 'godlike' form, a man and a lion has been combined to embody the attributes associated with each of the combined concepts (Dalton 2003). This kind of symbolic attribution is not an ancient technique reserved for history books and museums. Comic book superheroes are just one of many modern-day examples of the same phenomenon where concepts are combined for a new concept to emerge that inherit attributes from both input concepts (Hedblom et al. 2018).

The underlying phenomenon in these combinatorial examples is an analogical transfer in which attributes are transferred from one source domain to a target domain (Lakoff and Johnson 1980). Through the use of metaphors and representative attribution, art is a form of communication that has been prevalent throughout artistic creations until modern times. As an example, consider the metaphorical, and often politically charged, images by street artist Banksy. In their famous art piece Flower Thrower, a weapon (likely a brick or even a grenade) has been replaced with a flower bouquet, thus, transferring the notion of peace and love associated with flowers into the aggressive stance and clothing of the thrower. Exactly how we are supposed to interpret the image is not defined, but the universality of the conceptual spaces of the analogical transfer evokes common ground between the viewers.

Symbolism and conceptual metaphors are common tools in artistic expression, and many even go so far as to claim that art is all about sending a message. Balter (2008, p. 709) expressed it *as* "...art is an aesthetic expression of something more fundamental: the cognitive ability to construct symbols that communicate meaning...".

However, not everyone agrees that artistic expression is about communication or representation. While some artists are actively trying to tell a story with their work, the *l'art pour l'art* movement argues that art is in its own right enough. There needs to be no intentionality behind brush strokes nor representative attribution to color choices. Instead, just having *form* is enough for something to be art.

As Bender and Koller (2020) argued that generative AI created form, an interesting perspective emerge when speaking about computational creativity. As a prominent contributor to the field of computational humor and linguistic creativity, Veale (2014) described computational creativity (CC) as "...the scientific study of the creative potential of computers". As such, CC is a branch of AI research that stretches the full range of creative abilities: problem-solving, music composition, narrative and text generation, evaluation of CC systems and, naturally, computer-generated art.

*(a) Vincent van Gogh* *(b) Andy Warhol* *(c) Sandro Botticelli* *(d) Salvador Dali* *(e) Pablo Picasso*

*Figure 2: Midjourney generated images with the prompt "A dog in the style of ...". The images are selected from the first generated batch and not modified in any way.*

There are numerous examples of CC system that aims at creating art. One of the most famous examples is the robot and expert system AARON (Cohen 1995). Designed by creator Harold Cohen, AARON was one of the longest consistently running computational art projects. During the span of four decades, AARON's art style developed into increasingly complex and sophisticated forms as a consequence of the added information about the world. Prompted with sentences, AARON physically painted remarkable drawings and paintings more or less entirely autonomously. Many of the paintings were sold and showcased in different museums.

The Painting Fool is another influential art project that made portraits of the interactive audience (Colton 2012). Focusing less on the creative process and more on the creative agent, the Painting Fool was given a simplified type of 'personality'. Searching the Internet for news articles, it was given a mood that reflected the content of the articles which, in turn, influenced its choice of painting style. It could also opt out from performing the portrait if it 'did not feel like it'.

Today's perhaps most famous art generation system, Midjourney, is able to do an exceptionally wide variety of art pieces in relationship to the prompts that it receives. Based on using deep learning, transformer models and stable diffusion techniques to 'intelligently reproduce' images, its models are trained on millions of images on particular subjects or art styles. To demonstrate the power of today's tools, Figure 2 displays a comparative study of some of Midjourney's visual range by asking it to reproduce famous artists' painting styles based on the prompt "a dog in the style of ...". The generated images are rather visually extraordinary and while the artists might not have created exactly these images, the characteristic touch of the original artists shines through.

## 3. DISSECTING CREATIVITY: PRODUCTS, PROCESSES AND AGENTS

Taking one step back from art as a communicative expression and looking at creativity as a research area, one of the main problems is that the term itself lacks any clear definition. As Veale (2012, p. 1)) rightly pointed out "As soon as we think we've hemmed it in with a tight, rule-based definition, creativity is already hard at work on an escape plan". Much like with intelligence, creativity seems to expand in scope as we start to critically analyze it. Despite this complexity, the phenomenon of creativity is a fundamental component of both our society and our individual intelligence, and as such deserves to be investigated with the full range of scientific methodologies available.

What is often referred to as the 'the standard definition of creativity' states that a productor idea is required to be both novel and valuable/effective for it to be considered creative[8] (Runco and Jaeger 2012). While this makes it easier to look at objects and ideas from a creative perspective, very rarely is creativity an attribute exclusive to an object. While a painting may be creative, so too can the process by which the painting is made, and it is even more common to speak of the painter as the creative one.

To analyze creative ability, Rhodes (1961) proposed four dimensions worth investigation: persons, process, press and products. Commonly known as the four Ps of creativity, the one that stands out the most is 'press'. It is described as the creativity that emerges due to the relationship between humans and their environment, essentially the cultural context. While these were interpreted from evaluating the creative potential in humans, and not so much coring out the requirement of creative ability, three of these will be used as a foundation for analyzing the creative potential in computers: The creative product, the creative process and the creative agent, the fourth (press) will be embedded into the analysis of the other.

### 3.1 The Creative Product: The What

In the standard definition of creativity, the main criterion for something to be deemed creative was that the product was novel and valuable/effective (Boden 2009; Runco and Jaeger 2012; Stein 1953). This provides a straightforward methodology to evaluate the creative character of many products. Art pieces are creative if they portray something novel and people want to purchase them or view them in a museum. Likewise, ideas are creative if they are novel and effective in solving the problem.

#### *3.1.1 Novelty*

Looking at novelty, several researchers have distinguished between the societal impact of said novelty by separating little *c*-creativity vs. big *C*-creativity (Simonton 2013). Similarly, in the setting of CC Boden et al. (2004) distinguished between *P*-creativity and *H*-creativity. *P*-creativity, or *c*-creativity, is the personal discovery or the creation of something that is novel to the creative agent. Humans frequently do this, and without much thought, but it is perhaps most prominent in children that regularly create and discovers new ideas. Few children's drawings end up in a museum, and while the made-up words of children are creative, they might not be particularly useful in the grand scheme of things. Likewise, the generated art by a hobby artist represents a novel product to the agent, but it might not sell for a great price or be of interest to anyone but the creator themselves. Therefore, for the truly revolutionary creative, Boden et al. (2004) introduced *H*-creativity, or *C*-creativity, to highlight the creative products and ideas that go down in history, for instance, scientific discoveries, useful inventions and paradigm shifts in artistic techniques.

While important for all forms of creativity as stated by the standard definition, in *H*-creativity it is particularly pivotal for creative products to have value.

#### *3.1.2 Value/effectiveness*

Compared to novelty, value/effectiveness is often entirely disjoint from the product itself and instead exists as a relationship to the context in which the product is placed and the people who benefit from it. Take a chair that is designed without a seat, it might be a novel product, but a chair that does not afford[9] sitting, is useless. This means that if the 'meaning' of a product is not realised it is rarely considered creative.

For artistic domains, the value of art is the result of an opaque combination of complexity and skill of the artist, fame and exposure of the art piece and, for lack of a better word, general consensus—essentially Rhodes's (1961) *press*. As an example, consider Da Vinci's Mona Lisa, in many ways, it is neither innovative in its subject, as there are numerous historical portraits, nor is it particularly innovative in its painting technique, size or color combination. Despite this, it is considered the world's most valuable piece of art and millions of people travel to Paris for a chance to see her.

Smith (2005) argues that as value exists 'outside' the cognitive sphere it does not help provide any clarity to what should be deemed creative. Taking this into account, Smith and Smith (2017) proposed a 1.5 criterion model for creativity by highlighting that it is enough that the product has *potential* to be valuable.

However, it is not only the creative product that is subjugated to the complexity of attributing value. Klausen (2010, p. 249) suggests that is instead "... preferable to speak of a process which has a propensity for resulting in a novel work".

### 3.2 The Creative Process: The How

One prominent perspective of creativity is that it is a series of cognitive processes directly related to problem solving (Newell et al. 1962). While this might make sense from an evolutionary perspective in the sense of advancing society through inventions and scientific discoveries or even attracting mates through some emotionally charged 'music' outlet, unless one interprets the purpose of art as a form of representative communication, it feels unintuitive to position artistic creativity in the area of problem-solving.

Therefore, instead of looking at the motivations for the processes involved in creative behavior, it might be more useful to speak of the nature of the creative process. While there are overlaps and uncertain borders to how such processes might take place, we turn to Boden's (1998) three types of creativity: *Combinational, explorative* and *transformation*, as a categorical starting point.

#### *3.2.1 Combinational Creativity*

One of the most influential theories of combinatorial creativity is the theory of Conceptual Blending (Fauconnier and Turner 1998). In conceptual blending, two conceptual spaces are merged into a novel blended space by connecting the features and attributes in the two domains following the principles of analogical reasoning. The new concept inherits properties from each domain while emergent properties and optimization rules ensure that the blend turns into a coherent concept.

The mentioned Palaeolithic statue The Lion Man is an example of a conceptual blend. Another real-world example is a 'houseboat' which embodies the affordances and properties of both houses and boats. However, many compound word combinations display conceptual blending and creative behavior (Righetti et al. 2021). Important to note is that conceptual combinations are not exclusively based on physical properties. As the blend is based on conceptual spaces, any abstract and associative elements of the conceptual spaces can also be mapped and inherited during the merge. In that way, conceptual blending is assumed to be the cognitive foundation for the generation of novel concepts through combining what is already known in new ways.

### *3.2.2 Exploratory Creativity*

One aspect that combinatorial theories of creativity only implicitly deal with, or deal with in a theoretical space, is the phenomenon in which the effectiveness and evaluation of the novel concepts are generated. To understand what this means on a cognitive level, one original definition of creativity argued that creativity is an associate process in which "...associative elements are combined into new constellations..." (Mednick 1962, p. 221). While based on combinatorial creativity, the main point of The Associates Theory of creativity is that creativity is a process in which the mind flows freely and often through analogical reasoning makes new connections between conceptual elements in the conceptual spaces. However, exclusively making associations are not enough to evaluate creative ideas. To deal with this, research demonstrates that in the creative process as a whole, associative thinking is combined with analytical thinking (Groborz and Necka 2003). This means that the creative process acts as a cyclic process combining divergent and convergent thinking (Gabora 2010), allowing for both associations to be made, while also focusing and analyzing the results. This cyclic relationship has been captured in several theories aiming to explain the creative process (e.g. Wallas' model of creative thought (Wallas 1926) and the Geneplore model (Finke et al. 1996)).

In this cyclic process, exploratory creativity is the creative process in which conceptual spaces and domains are manipulated in such a way that novel products and content arise. As an example, artists who conform to a particular conventional technique yet still create novel pieces of art perform a sort of exploratory form of creativity.

### *3.2.3 Transformational Creativity*

Similar to exploratory creativity, transformational creativity is about searching conceptual spaces but while doing so breaking the bounds of the particular conceptual domain they are in and expanding beyond the limits of that particular conceptual space. As an example, consider the introduction of moving pictures as the inclusion of a temporal dimension to the conceptual domain of photography. Like this, many cases of successful transformational creativity become by definition *H*-creativity as they transform the paradigm of the conceptual space.

Another example is how Jackson Pollock introduced a completely different art technique with his action painting. Supposedly inspired by the sand painting technique by the Navajo in which colored sand is poured from the hands into intricate patterns and pictures (Staker 2019), Pollock let paint drip and splash from a brush in seemingly random chaos. The results were vastly different, but using a technique in which the paint is added to the target without physical contact was the same, demonstrating the exploratory process. Combined, the results together with the technique were new not only to Pollock but to the whole world, showcasing the transformational character of the creative process. To expand beyond the borders of already known conceptual spaces requires a much larger knowledge foundation than what many people actually possess.

One consequence of this is that many cases of transformational creativity are created through serendipitous happenstance and were never truly intended at all. As this is related to the intent of the creative agent, this will be elaborated on in the upcoming section.

### 3.3 The Creative Agent: The Why

One of the original motivations to study creativity was to investigate extraordinarily creative individuals in the sciences or the arts and in so doing, try to establish a creativity quota (cf. intelligence quota). This is commonly referred to as the Creative People Approach (DiPaola and Gabora 2007) and several tests were designed to determine the degree of creative thinking an individual possessed, the perhaps most famous is Torrance's (1966) test for creative thinking (TTCT).

While the idea of a creative quota has largely been left in the history books, due to the development of computational methods that simulate creative performance, it is worth considering what would be appropriate methods to evaluate the creative ability of generative AI and other creative computer systems.

When evaluating creative agents, things that need to be considered are related to two main categories: *Performance* and *motivation*.

#### *3.3.1 Performance*

The Threshold Theory argues that an agent needs to have intelligence that exceeds a particular threshold in order to display creative ability (Kozbelt et al. 2010). This makes sense as a baseline is required in order to successfully perform both 'diverse enough' divergent thinking and 'accurate and sophisticated enough' convergent thinking. In fact, up to a certain point there is a correlation between creative ability and level of intelligence (Preckel et al. 2006). Despite this, school children display a decrease in creative thinking as they grow in cognitive maturity likely due to more routine behavior being applied in everyday life (Kim 2011). Regardless of the reason, it showcases that the relationship between intelligence and creative ability is not linear.

Another complexity to this is the relationship any application of intelligence has with knowledge. Schank and Cleary (1995, p. 229) suggest that creativity is the "...intelligent misuse of knowledge structures...". This requires the agent that is performing the creative process to be in possession of enough knowledge that can be misused. In order to be able to engage in any of the three types of creativity, the involved conceptual spaces need to be rich enough to allow for combinations, modifications and explorations. However, this is particularly important for *H*-creativity which requires the creative products to be outside of the scope of what is universally known.

The final component of performance is skill. In order to be able to participate in the creative process, the agent is required to have the necessary motor and cognitive skills to be successful. For humans, this means spending lifetimes learning and practicing particular skills in which they grow in expertise and knowledge.

#### *3.3.2 Motivation*

The system behind the Richter-Window was not considered creative as it had no intent to do something creative. Klausen (2010) explained this by arguing that a minimum of an "intention of novelty" is required for someone/something to be creative.

There are many reasons why an agent chooses to engage in creative tasks. Most of these reflect back to either the creative product, as in the purpose of developing new ideas or products; or to the creative process, with the motivation that creative and artistic endeavors are fun and entertaining and, thus, meaningful as a process. However, people also have motivations that redirect back to the agent themselves. Consider things like mandalas and the generation of sand patterns in Zen gardens. Such 'art' is not created for the purpose of generating a product but as a form of spiritual meditation intended to transform the creative agent. Further emphasizing this perspective, art is a common tool in therapy and studies have shown that doing art reduces stress (Bolwerk et al. 2014). In agent-centric creative motivation, the resulting product is of no importance whatsoever, nor is the process required to take any particular form or to be particularly innovative and creative. Proponents of this perspective, argue that creativity by definition, happens internally as part of the cognition of an agent, and not through any external process or evaluation of products (Sen and Sharma 2011).

To complicate things, many creative ideas, especially transformational ones, are the result of serendipity, also referred to as 'accidental creativity.' One supposed example of a serendipitous creative product of *H*-creativity level is the chocolate-chip cookie. The legend goes that bits of chocolate accidentally fell into the cookie batter and, impossible to take out, was simply baked into the cookies. Accounting for these situations, Finke (1996) argues that there is an 'illusion of intentionality' in creative processes and that agents often do not intend to be creative. This would mean that it is in retrospect that both motivation and intentions are added to problem-solving and artistic processes.

In this section, some of the main criteria of creativity theory were outlined. To evaluate the state of the art in generative AI on these metrics, the next section presents an evaluative study on the creative ability of Midjourney.

## 4. A STUDY IN GENERATING DOGS: A CREATIVITY ANALYSIS OF GENERATIVE AI

In order to evaluate the creative ability of generative AI, Midjourney was selected as the main running example. This selection was partly due to convenience sampling as it is commercially freely available and accessible, but also due to the rather remarkable success the tool has seen on the art scene. As many of these systems are based on similar underlying AI technologies and/or have implementations that are trademarked, using Midjourney is argued to be sufficient for an explorative evaluation of the creative ability of such systems.

The selected methodology is a brief, yet systematically effective, study performed by prompting Midjourney with the task of 'generating dogs' by, through different prompt formats, targeting the core components of the criteria mentioned in the previous section. Each presented figure used in the argumentation, corresponds to the very first generated image based on the documented prompt and has not been edited or manipulated in any way. As Midjourney by default creates four variants for each prompt—presented as one image, in the cases where only one illustration is presented, the version that was deemed the 'most representative' was selected.

The presented analysis will touch upon all the relevant criteria by asking the questions: *Is Generative AI Producing Creative Products? Is Generative AI a Creative Process?* and *Is Generative AI a Creative Agent?*

### 4.1 Is Generative AI Producing Creative Products?

For a product to be considered creative two components were required: Novelty and value/effectiveness.

*Figure 3: Midjourney generated images with the prompt "a dog in the style of Frida Kahlo".*

For computer generation, novelty is exceptionally easy to accomplish as even a random function can produce an exponential number of outcomes in relation to the number of included variables. The Richter-window consists of 11.263 squares that can take any of 72 colors leading to $72^{11.263}$ or $2.6708 \times 10^{19}$ unique windows.[10] This means that while the window is auto-generated with very few parameters, essentially only the different colors, the window is one of an incomprehensibly large number of possible outcomes. Each one, from a mathematical point of view, would be equally novel and could be argued to be a form of simplified P-creativity, much like repetitive drawings. Yet, few would probably argue that this would display any higher level of creative variety.

Similar reasoning can be attributed to the products made by generative AI where it is possible to reduce these to describe pixel grids in different colors. Any number of 'dog' images can be created as a relationship between the training data and the degree of freedom in the algorithms underlying the generative models. In Figure 2, one out of four variants of each artist were presented, and Midjourney could keep producing dog illustrations of the same artists indefinitely as the variables are not as restricted as with the Richter-Window. To test this hypothesis, Figure 3 presents three generated image collections with four variants each of the prompt "a dog in the style of Frida Kahlo." While each illustration is unique, the variation of the generated illustrations is rather limited.

However, the main difference to the Richter-Window is the ability to generate shapes with a 'meaningful' connection to the words in the prompts. The 'Frida Kahlo dogs' are all really lovely as hyper-realistic paintings that have high accuracy in relation to the intended referential attribution. Granted, they look more like attributes associated with pictures of Kahlo rather than her art style (e.g. flower crowns), but this might be a problem with the specificity of the prompt, rather than Midjourney's skill in determining the painting style of Kahlo.

Ascertaining the value of these autogenerated images is significantly more complex. Value of art is based on aesthetics within the presenting context, and as the famous philosopher David Hume correctly ascertained "Beauty is no quality in things themselves; it exists merely in the mind which contemplates them; and each mind perceives a different beauty" (Hume 2005, p. 35). There is therefore no standard by which one can measure the success of artistic products' values other than the way in which they are greeted in the sociocultural setting they are presented in.

Already decades ago, images from the computational painter AARON (Cohen 1995) were on display in museums and recently an AI-generated image won first prize in an art competition. With these truths mind, it is hard to argue that images generated by AI do not live up to the criteria of being creative products.

### 4.2 Is Generative AI a Creative Process?

As we have established that generative AI can produce creative products, the process by which it does this is required in-depth analysis.

Many people are familiar with the Imitation Game, or the Turing test, in which a human evaluator is asked to judge the intelligence of a computer by evaluating whether it is communicating with a human or an artificial agent (Turing 1950). Less familiar is the Lovelace test, in which a computer agent is only considered intelligent if it can create something original, essentially being creative (Bringsjord et al. 2003).

The test was designed based on Lovelace's skeptical perspective on machine creativity. She writes "The Analytical Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform" (Lovelace 1843).

In truth, it is quite hard to instruct generative AI to do what one wants. As these models act like blackboxes without instructions and explainability, prompting the generative AI has almost become an art form in and of itself. This lack of explainability is a big problem for predictive models in more ethically charged areas and a lot of work is being done on improving the transparency in this research field (Lipton 2018). Arguably this is less of a problem for computational artists where also the human creative process is largely unknown. However, it does mean that these models behave exactly like the system in the Chinese Room argument. Only, compared to programming code such as that behind the Richter-Window, the level of competence by which these systems perform has given them a different status in the debate about the level of intelligence and, by consequence the level of creativity, they possess.

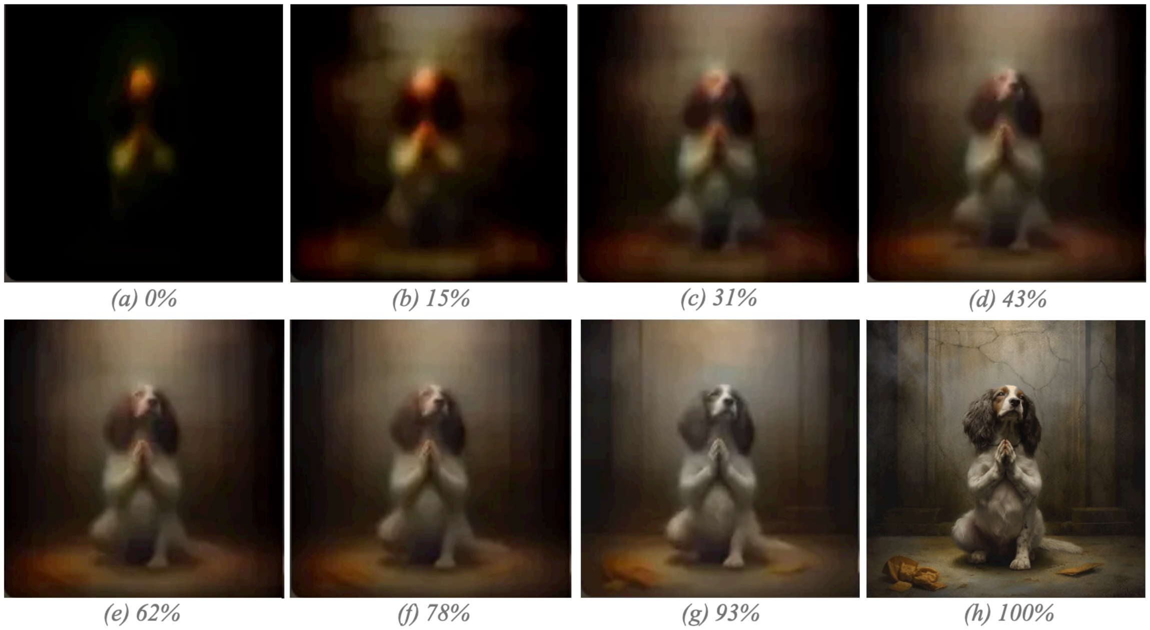


*Figure 4: Midjourney's generative process of creating the prompt "a praying dog."*

Focusing exclusively on the creative process, it was proposed to take place as three different types. For style transfer from one domain to another like that showcased in Figure 2, a combination of both combinatorial and exploratory creativity can be considered to be at play. The system explores the conceptual spaces of dogs and the style of the artists and outputs a somewhat convincing combination. Much more remarkable is the symbolic attribution in the images. For instance, the 'Dali dog' presents surrealistic imagery with a watch on a top hat, what looks like a metronome (which I assume is a reference some of Dali's sculptures) and some creature-like things falling from the sky. Likewise, in Figure 3, essentially all dogs feel associated with what a conceptual space of Frida Kahlo could look like based on the flower crowns, clothing and embellishments that (to an untrained eye such as mine) look Mexican-inspired, not to mention the rather solemn facial expression on the dogs. As this output is exclusively (as far as we know of Midjourney's implementation) due to the model having learned form, these associative elements emerge through an almost analogy-based exploratory creativity.

Simultaneously, as most prominently demonstrated by the 'Kahlo dogs,' there is not much in terms of transformational creativity. The system remains within its allotted boundary and the basis generated from its training data.

In addition to these processes being subjugated to the black-box phenomenon, the programmed process that the artistic system undergoes from prompt to generated image tends to be trademarked. This means, that it is essentially impossible to take a critical look at the 'cognitive' aspects of the process. What can be done, is to follow the generative process. In Figure 4, Midjourney's transition from idea form to the final product is displayed as intermittent steps denoted with completion percentages. The images do reveal that the system is quite clear on what it wants to achieve in terms of pose and visual concept quite early on (perhaps even before the visual generation starts) and then fine-tunes the image by polishing details, edges and structures as a result of what is likely diffusion models. In many ways, this approach is a similar methodology to that of many artists that draw outlines of their intended object only to zoom in to add details and final touches as the picture emerges.

Generative AI might not (yet) be able to break the boundaries of their training data and alone venture into what would be transformational creativity worthy of the epithet of *H*-creativity. However, it is also not fair to claim that the process underlying these tools differs so greatly from human creativity, which cognitive processes remain largely unknown, that they cannot be deemed to express, at the very least, little *c*-creativity.

### 4.3 Is Generative AI a Creative Agent?

It is when generative AI is being evaluated as creative agents, that the real complexities arise. Even with the most liberal interpretation of the requirements of agency, generative AI tools fall short on some critical accounts.

The Threshold Theory of creativity required a certain level of intelligence for a system to be creative (Kozbelt et al. 2010). Exactly how generative AI systems should be evaluated based on their intelligence quota is of course outside the scope of this article. However, a case can be made that they are indeed intelligent enough. Perhaps it is not the same general intelligence found in humans, but they do possess the required intelligence in the sense of completing cognitive tasks.

*(a) "a dog with typical dog attributes"* *(b) "a dog without dogness"* *(c) "a fake dog"*

*Figure 5: Midjourney generated images with the mentioned prompts to test the understanding of the conceptual space of 'dog'.*

For this, they require enough knowledge. Humans spend lifetimes learning and perfecting their art through iterative experience and practice. Obviously, artists inspire one another and there is a certain 'creative theft' within artistic circles. In comparison, generative AI is instead force-fed millions of art pieces without perceptual, or experiential grounding of these pieces into a real-world situation of personal experience. An agent presented with millions of pictures of dogs, surely knows what a dog looks like, but they will have no understanding of what a dog really is.

Testing this, Figure 5 contains the result of three different prompts designed to question Midjourney's conceptual understanding of what dogs are. In Figure 5a, the tool was asked to produce a visual response to the prompt "a dog with typical dog attributes". Two of the dogs reveal the lack of understanding of what typical dog attributes are as what appears as a medieval-inspired armor and a steampunk 'Christmas tree' outfit were deemed appropriate results to the prompt. Likewise, the results of the exclusion prompt "a dog without dogness" presented in Figure 5b, showcase that the tool does not understand what to remove as the result all look very much like prototypical dogs. The lack of success with the prompts in Figure 5a and 5b is likely the result of not having enough conceptual understanding of all that a dog 'is'. Figure 5c shows a much more successful interpretation of the prompt. Asked to generate "a fake dog", three (upper-right being the exception) of the four dogs appear to be generated as fake dogs in the format of rather nightmarish statues and toys. That this prompt was more successful can be a result of more combinatorial creativity being at play as the conceptual spaces of 'fake' and 'dog' remain conceptually intact and simply combined as a linguistic adjective phrase.

If having a level of conceptual understanding truly matters when it comes to generating art is, I guess, a question for the evaluator of such a dog painting by in retrospect choosing to assign creative intent or not (Finke 1996). However, it does prevent the possibility to interpret the art piece as a form of communication as the model has nothing to 'say'.

In the role of a human evaluator of the generated dog images, it is quite clear that the skill by which Midjourney performs greatly surpasses that of most people. The quality of the pictures of Midjourney, Stable Diffusion and DALL-E all exceed those of laymen and experts alike, both in terms of the depth of quality as well as the width of styles. There might be things they cannot do, like expanding beyond their own programming or training data, but most human artists cannot do that either.

It is worth mentioning that as humans, we spend a great deal on the emotional connection to our experiences, and skill and training are tightly connected to internal reward systems. In Colton (2008) argued that a core component of a computational system to be deemed creative is having the ability to appreciate what has been created. The song-writing musician Nick Cave responded to a chatGTP-generated lyrics "in the style of Nick Cave" by saying it was "...a grotesque mockery of what it is to be human..."[11] Exactly, what it means to be human is of course a philosophical question of a whole different dimension, but the creative spirit that is expressed within these tools displays a simulation, not a genuine reflection during the process nor an appreciation of the created products.

Our emotional capacity, will to live, personal pursuits and persistence when faced with challenges do not only translate into our artistic expression but also into our motivation for creation. This relates to a crucial component of the artificial general intelligence debate, namely whether an AI agent would be able to develop a 'free will' of sorts in which it can develop motivation and intentions for its choices and actions. While the notion of free will is a rather complex term, any system bound to perform an action requires an initial incentive. Landgrebe and Smith (2022) argue against that motivational drive could ever be mathematically simulated to the appropriate level for a computational system to display a 'will'. Even disregarding the work on reinforcement learning and curiosity-driven learning (see Burda et al. 2018), modelling will is also possible to with remarkably simple systems such as those demonstrated with Braitenberg Vehicles[12] demonstrate a behavioral response to stimuli. Similarly, Landgrebe and Smith's (2022, p. 134) argument on how no one would request a sentience analysis of a steam engine, it still displays appropriate cause-and-effect behavior as a response to stimuli. In both examples, what is demonstrated is the same as within The Chinese Room argument, namely that a system requires quite little in terms of input and transformation rules to still initiate actions and behavior that makes sense.

For Midjourney, the incentive comes from a human prompt and the process is defined by training data and the implementation of the system to generate the art pieces. Having a type of external 'button-presser' is hardly something that we think of as a required component for human artists where incentives and persistent motivation can independently drive the whole creative process. This means that no matter how one twists and turns the definitions, one thing that the current states of generative AI do not possess is motivation and as such they lack the autonomous agency required to be deemed creative agents. Professional artists often produce art pieces purely based on customers' requests, just like a prompted generative AI system. However, much like The Painting Fool (Colton 2012), a human artist also has the freedom to refuse or to paint something else. Naturally, it is always possible to return to the 'illusion of intentionality' argument (Finke 1996) and claim that any artistic product that is deemed to be creative in retrospect must require a creative agent as well. However, the problem remains as this evaluation can only be performed by an external perceiver and not by some convergent process of the system itself.

Ultimately, even in the setting of serendipitous happenstance, the attribution of creativity to a computational system requires the human evaluator and the context to be taken into account.

### 4.4 Stronger Together: Co-Creative Systems

Taking a comparative look at art generation between human actors and AI systems, fundamental differences appear. Humans create art for a range of reasons like product development, enjoyment and even therapy. AI does it because we tell it to.

What was left unsaid for the sake of argumentation, is that the honest truth about both the Richter-Window and the award-winning art piece Théâtre D'opéra Spatial is that neither was created in isolation by a computational system. The Richter-window's framing in terms of the selected 72 colors, the grid pattern and the placement in a church, was designed by Richter, not the computer program - and it is here that the truly creative design lies, not in the different variants of color combinations. Likewise, Midjourney did not alone achieve the masterpiece that won the art competition. Instead, the image was produced as a collaborative effort between Jason Allen and the system. Allen prompted the system with a complex text description and as Midjourney produced the initial image, Allen reported that he iteratively fine-tuned and re-prompted the design of the image until he was satisfied (Allen 2022). While Allen probably did not have this exact picture in mind, and probably would not have been able to create this precise version of it, the winning image was never really constructed by a generative AI at all. It was a co-creative process in which a human was instructing a computer tool based on his own aesthetic preferences.

While it is tempting to think of this as a reduction in the creativity of the system, the main point is not that the system is not creative enough, it is that it does not appreciate the art piece enough. Midjourney is not able to say "Enough! This is the perfect image," or "I like this one!" Instead, it is possible to prompt Midjourney to morph and manipulate the images indefinitely. As long as it does not perform an evaluation based on personal experience and developed preference, Midjourney will never be satisfied. Satisfaction and dissatisfaction alike are entirely the responsibility of the human user.

## 5. THE LIGHT AT THE END OF THE TUNNEL

AI-generated art is beautiful, novel and truly remarkable as a technological feat. Tools like Midjourney, Stable Diffusion and DALL-E have more realistically put humanity in her place not only as an artist but as an appreciator of art. We are no longer alone in creating artistic forms that can touch us both intellectually and emotionally. Together with these tools, we are offered the assistance of increased production speed, enhanced skills and an increased chance for serendipitous innovation. Despite this, we should not forget that these tools do not possess any intentionality, there are no conscious efforts, no aesthetic evaluation based on a lifetime of experience and there is no emotional elevation to particular generated images. These systems are unable to appreciate the art they generate.

This means, as of yet (and perhaps forever), these tools are still just that: tools. They do not perform based on any level of self-motivation, have neither conscious awareness (as far as we can claim to be able to measure) of the process nor any active decision-making throughout the creative process. Instead, they are trained on human-generated art, prompted by human intent and, in a co-creative process, asked to improve and modify their output based on a human user's preferences and expectations. Ultimately, these generative AI systems fall short with respect to the *"creative potential"* that Veale (2014) pointed out was the main study in computational creativity as it is a human user that guides the creative process and a human context that evaluates the value of the produced content.

It is clear that systems such as Midjourney will not rule the (art) world. Simultaneously, as their competence, generalizability, speed and efficiency far exceed that of human artists, I think it is safe to say that while they may not *rule* the world, they will without a doubt *run* the world.

## NOTES

1 Intentionality here is intended in its philosophical interpretation as the mind's ability to form representations and should not be confused with the ability for having intentions.
2 https://openai.com/gpt-4, accessed: 2023-06-08.
3 Synthetic phenomenology refers to the phenomenal states that any synthetic system (e.g. a robot or artificial agent) possess.
4 https://openai.com/product/dall-e-2.
5 https://stablediffusionweb.com/
6 https://www.midjourney.com/
7 Here understanding is thought to represent intentionality as the relationship between a symbolic representation (such as a word, an icon or a piece of representative art) and the intended referent 'in the real world.'
8 A wide variety of synonyms to these terms have been used in the history of creativity research, but they all capture essentially the same notions.
9 Affordances in the sense of Gibson (1977).
10 Putting this into perspective, an estimate is that each year 2.25 x $10^{19}$ grains of rice are produced world-wide.
11 https://www.theredhandfiles.com/chat-gpt-what-do-you-think/
12 Braitenberg Vehicles are simple programmed agents that respond to stimuli in different ways (e.g. moving towards/away from light). Braitenberg Vehicles were introduced as a thought experiment to demonstrate how it is possible to generate 'biological' behaviors with very little input (Braitenberg 1986).

## REFERENCES


Allen. J. 2022. *How Jason Allen used AI to win Art Prize*. Interview on: Disruption Theory. https: //www.youtube.com/watch?v=yAQz7NO2E_U

Balter. M. 2007. On the origins of art and symbolism. *Science*, 323(5915):709-711, 2008. doi: 10.1126/science.323. 5915.709.

Belz. C. 2007. Das Richter Fenster, Documentary.

Bender E. M. and A. Koller. 2020. Climbing towards NLU: On meaning, form, and understanding in the age of data. In: *Proceedings of the 58th annual meeting of the association for computational linguistics*, pp. 5185-5198. https://aclanthology.org/2020.acl-main.463.pdf

Boden. M. A. 1998. *Creativity and artificial intelligence.* Artificial intelligence, 103(1-2):347-356.

________. 2009. Computer models of creativity. *AI Magazine*, 30(3):23-23.

Boden M. A. et al. 2004. *The creative mind: Myths and mechanisms.* London: Psychology Press.

Bolwerk, A., J. Mack-Andrick, F. R. Lang, A. Dörfler and C. Maihöfner. 2014. How art changes your brain: Differential effects of visual art production and cognitive art evaluation on functional brain connectivity. *PloS one*, 9(7): e101035, 2014.

Braitenberg. V. 1986. *Vehicles: Experiments in synthetic psychology.* Cambridge, MA: MIT Press.

Bringsjord, S., P. Bello and D. Ferrucci. 2003. Creativity, the Turing test, and the (better) Lovelace test. In: *The Turing test: the elusive standard of artificial intelligence*, pp. 215-239. New York: Kluwer.

Burda, Y., H. Edwards, D. Pathak, A. Storkey, T. Darrell and A. A. Efros. 2018. Large-scale study of curiosity-driven learning. arXiv preprint arXiv:1808.04355

Cohen. H. 1995. The further exploits of AARON, painter. *Stanford Humanities Review*, 4(2):141-158.

Colton. S. 2008. Creativity Versus the Perception of Creativity in Computational Systems. In: *AAAI spring symposium: creative intelligent systems*, vol. 8:7.

________. 2012. The Painting Fool: Stories from building an automated painter. *Computers and Creativity*, pp. 3-38.

Colton S. and G. A. Wiggins. 2012. Computational creativity: The final frontier? In: *ECAI*, vol. 12:21-26. Montpelier. doi: 10.3233/978-1-61499-098-7-21.

Dalton. R. 2003. Lion man takes pride of place as oldest statue. *Nature*, 425(6953):7-8.

Devlin, J., M.-W. Chang, K. Le and K. Toutanova. 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv* preprint arXiv:1810.04805.

DiPaola, S. R. and L. Gabora. 2007. Incorporating characteristics of human creativity into an evolutionary art algorithm. In: *Proceedings of the 9th annual conference companion on Genetic and evolutionary computation*, pp. 2450-2456. New York: Publisher: Association for Computing Machinery.

Fauconnier G. and M. Turner. 1998. Conceptual integration networks. *Cognitive Science*, 22(2):133-187.

Finke, R. A. 1996. Imagery, creativity, and emergent structure. *Consciousness and cognition*, 5(3):381-393.

Finke, R. A., T. B. Ward and S. M. Smith. 1996. *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Gabora. L. 2010. Revenge of the "neurds": Characterizing creative thought in terms of the structure and dynamics of memory. *Creativity Research Journal*, 22(1):1-13.

Gibson, J. J. 1977. The theory of affordances. In: R. Shaw and J. Bransford (eds.). *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, pp. 67-82. Hillsdale: Lawrence Erlbaum.

Groborz M. and E. Necka. 2003. Creativity and cognitive control: Explorations of generation and evaluation skills. *Creativity Research Journal*, 15(2-3):183-197.

Hampton, J. 1993. Prototype models of concept representation. In: I. V. M. et al. (ed.), *Categories and concepts: Theoretical views and inductive data analysis*, Cognitive science series, pp. 67-95. Cambridge: Academic Press.

Hedblom, M. M., R. Penaloza, O. Kutz, and G. Guizzardi. 2018. Under the super-suit: What superheroes can reveal about inherited properties in conceptual blending. In: *Proceedings of the Ninth International Conference on Computational Creativity*, pp. 216-223. ICCC, Salamanca, 25-29 June. Association for Computational Creativity (ACC).

Hume, D. 2005. *On Suicide*. Harmondsworth: Penguin.

Kim, K. H. 2011. The creativity crisis: The decrease in creative thinking scores on the Torrance tests of creative thinking. *Creativity Research Journal*, 23(4):285-295.

Klausen, S. H. 2010. The notion of creativity revisited: A philosophical perspective on creativity research. *Creativity Research Journal*, 22(4):347-360.

Kozbelt, A., R. A. Beghetto and M. A. Runco. 2010. *Theories of creativity*. Cambridge: Cambridge University Press.

Lakoff G. and M. Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Landgrebe, J. and B. Smith. 2022. *Why machines will never rule the world: artificial intelligence without fear*. London and New York: Taylor & Francis.

Lipton, Z. C. 2018. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3):31-57.

Lovelace, A. A. 1843. Notes by A.A.L. In: *Scientific Memoirs Selected from the Transactions of Foreign Academies of Science and Learned Societies*, vol. III:666-731. London: Richard and John E. Taylor.

Manrique, H. M., C. J. Völter and J. Call. 2013. Repeated innovation in great apes. *Animal Behaviour*, 85(1):195-202.

Mednick, S. 1962. The associative basis of the creative process. *Psychological Review*, 69(3):220.

Metzinger, T. 2021. Artificial suffering: An argument for a global moratorium on synthetic phenomenology. *Journal of Artificial Intelligence and Consciousness*, 8(01):43-66.

Morriss-Kay, G. M. 2010. The evolution of human artistic creativity. *Journal of Anatomy*, 216(2):158-176.

Newell, A., J. C. Shaw and H. A. Simon. 1962/1958. The processes of creative thinking. In: *Contemporary Approaches to Creative Thinking*, University of Colorado. Presented at the aforementioned symposium. New York: Atherton Press.

Preckel, F., H. Holling and M. Wiese. 2006. Relationship of intelligence and creativity in gifted and non-gifted students: An investigation of threshold theory. *Personality and Individual Differences*, 40(1):159-170.

Rapaport, W. J. 2023. Is artificial general intelligence impossible. *Cosmos + Taxis* 12:5+6

Rhodes. M. 1961. An analysis of creativity. *The Phi Delta Kappan*, 42(7):305-310.

Righetti, G., D. Porello, N. Troquard, M. M. Hedblom, O. Kutz, and P. Galliani. 2021. *Asymmetric hybrids: Dialogues for computational concept combination*. In Formal Ontology in Information Systems: Proceedings of the Twelfth International Conference (FOIS 2021), vol. 344:81-96. IOS Press.

Runco, M. A. and G. J. Jaeger. 2012. The standard definition of creativity. *Creativity research journal*, 24(1):92-96.

Schank, R. C. and C. Cleary. 1995. Making machines creative. In: *The creative cognition approach*, pp. 229-247. Cambridge, MA: MIT Press.

Searle, J. R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3):417-424.

Sen, R. S. and N. Sharma. 2011. Through multiple lenses: Implicit theories of creativity among Indian children and adults. *The Journal of Creative Behavior*, 45(4):273-302.

Simonton, D. K. 2013. What is a creative idea? little-c versus big-c creativity. In: *Handbook of research on creativity*, pp. 69-83. Cheltenham: Elgar.

Smith, G. J. 2005. How should creativity be defined? *Creativity Research Journal*, 17(2-3):293-295.

Smith, J. K. and L. F. Smith. 2017. The nature of creativity: Mayflies, octopi, and the best bad idea we have. *Creative contradictions in education: Cross disciplinary paradoxes and perspectives*, pp. 21-35. Cham: Springer.

Staker, M. 2019. *Jackson Pollock and the Native-American Shaman*. PhD thesis, California State University Dominguez Hills.

Stein, M. I. 1953. Creativity and culture. *The Journal of Psychology*, 36(2):311-322.

Torrance, E. P. 1966. Torrance tests of creative thinking. *Educational and Psychological Measurement*. Princeton: Personnel Press.

Turing, A. 1950. Computing machinery and intelligence. *Mind*, 59(236):433.

Veale, T. 2012. Exploding the creativity myth: The computational foundations of linguistic creativity. London: A&C Black.

________. 2014. The revolution will be automated: Computers that create. https://www.slideshare.net/ kimveale/introducing-computational-creativity

Wallas, G. 1926. *The art of thought*, volume 10. San Diego: Harcourt, Brace.