

# AI and the Cluster Account of Art

Alice C Helliwell

In Ball, B., Helliwell A. C. & Rossi A. (eds.) 2024. *Wittgenstein and AI: Value and Governance*, London: Anthem Press.

Is AI art really art? This question has been the subject of much public discussion and is one that philosophical aesthetics should be well-placed to address. Unfortunately, there is no clear consensus within the discipline on how to tackle key definitional questions such as this. In the case of AI, we can add to this the unique challenge of works not made by humans. In this chapter, I argue for the utility of a Wittgensteinian approach to the question of whether AI art is art. This typically repudiates the need to provide necessary and sufficient conditions, when addressing the topic of AI art. Using Gaut's cluster account, I show that AI art can indeed count as art. I also demonstrate that the cluster account of art is particularly useful for thinking about art made by AI.

## THE CLUSTER ACCOUNT OF ART

The perceived failures of contemporary definitions of art (particularly a failure to garner any broad consensus amongst philosophers) led Berys Gaut to take a Wittgensteinian approach to art. Gaut was not the first to consider Wittgenstein's work in relation to the definition of art. Gaut's theory re-visits the work of philosophers in the 1950s who applied Wittgenstein's family resemblance approach to the question, 'what is art?', arguing for an anti-definitionalist approach (see Weitz 1956; Ziff 1953; Kennick 1958). These philosophers argue for two key points: first, that art cannot be defined (in terms of individually necessary and jointly sufficient conditions), and second, that art is a concept best characterised in terms of family resemblance (Gaut 2000). Instead of resemblance-to-paradigm as the model for the concept of art, however, Gaut turns to a 'cluster account' construal of family resemblance (Gaut 2000, 26). The cluster version of family resemblance that Gaut adopts comes from Wittgenstein's discussion of proper names and was further developed by Searle (1958). As Gaut writes,

A cluster account is true of a concept just in case there are properties whose instantiation by an object counts as a matter of conceptual necessity toward an object's falling under a concept... There are several [properties] criteria for a concept. (Gaut 2000, 26)

The way properties count towards the concept are as follows:

- 1) If *all* properties are instantiated in an object, then the concept applies to it. If *fewer* than all criteria are instantiated, this is sufficient for the application of the concept.
- 2) There is no one property which everything falling under the concept must have.
- 3) There are no individually necessary conditions for the application of the concept, but there are disjunctively necessary conditions. It must be true that some of the criteria apply if an object falls under the concept. (Gaut 2000, 27)

In the case of art then, there will be a list of properties whereby:

- i. If all properties are fulfilled by an object, it must be art, but if only some of the properties are fulfilled by the object, it can still be art.
- ii. There is not one property which all artworks will have.
- iii. For an object to be art, it must have at least some of the properties; it cannot have none of the properties, and still be art (though, in the case of art, Gaut states we might not yet know all of these properties). (Gaut 2000, 27)

Gaut's cluster concept of art is not a *true* cluster concept, however (even under the conditions that Gaut himself lays out). Gaut adds a necessary condition to his account, thereby violating condition 2). Gaut adds that given it is *artworks* in which we are interested, some action must have taken place at the genesis of something's being art:

An artwork is the product of an action, preeminently of a making (an artifact), or a performing (a performance). It is *artworks* that are involved here, since something is in each case done. Hence being the product of an action is the genus of the artwork and is thus a necessary condition for something's being art. (Gaut 2000, 29)

Gaut's account of 'making' or 'performing' here is permissive, as the account can still allow for found art; even found artworks are selected, and thus are the product of an action. This is sufficient under the cluster account of art, because "Selection adds to the range of properties that can be possessed by objects, and thus alters them, even if not physically." (Gaut 2000, 29).

Gaut argues that this modification to the cluster account does not compromise the nature of the account, as the necessary condition is minimal:

Being the product of an action is, however, a very thin generic condition, which does not distinguish artworks from any of the other products of action (philosophy, papers, chairs, pay freezes, angry words etc.) ... thus the modified cluster account holds that there is one necessary condition for something's being an artwork, but that is because of the notion of a work (the product of action) rather than because of the notion of art. (Gaut 2000, 29)<sup>1</sup>

With this minimal necessary condition in place Gaut proffers ten properties which may count towards something being a work of art (Gaut 2000, 28). Turning to AI, I will utilise these properties to evaluate AI works, arguing that they could meet several of the criterial properties, before returning to the necessary condition.

### **AI & THE CLUSTER ACCOUNT OF ART**

In order to evaluate AI works against the cluster account of art, I will begin with the list of ten properties in Gaut's (2000) account and propose how AI works could or could not meet them.

---

<sup>1</sup> Here I include these lengthy quotations because, as we shall see soon, the exact formulation of this necessary criterion will be particularly relevant when applying this account to AI.

### **i. possessing positive aesthetic qualities e.g., being beautiful, graceful, or elegant**

Possessing positive aesthetic qualities does not seem to be a barrier for AI works. There is no reason to think that AI images cannot be beautiful, even if you think the majority are not. Finding the work to have positive aesthetic qualities likely led to an AI work winning a prize at the Colorado State Fair (Roose 2022) and works made with AI have recently been described as “gorgeous” and “beautiful” in *The Guardian* (Jones 2022).

### **ii. being expressive of emotion**

Whether a work made by an AI could be expressive of emotion will come down to exactly what is meant by ‘being expressive’. Jenefer Robinson has provided us with an account of what it means for a work to have expressive qualities:

I would like to suggest that we should confine the term ‘expressive quality’ to those qualities in an artwork (or other things, such as merry brooks and anguished old oak trees) that are not only named by an emotion word but also arouse appropriate emotions. More particularly, *expressive qualities* are *qualities that can be grasped through the emotions that they arouse*. (Robinson 2005, 291-292)

If we adopt Robinson’s account of expressive qualities, it seems possible for an AI to produce works that have such qualities, i.e., works with qualities that can be named by an emotion word and arouse appropriate emotions. If an oak tree can have expressive qualities in this way, why not a work by an AI? What counts for Robinson is the reception of the work by the audience. She states that expressive qualities should “evoke corresponding emotions in audiences” which can alert the audience to the presence of said qualities (Robinson 2005, 292). We can already see these kinds of expressive qualities being described in works that are made with AI. For example, Jones (2022) describes the work of Gillian Wearing, who utilised DALL-E 2, saying: “You get a sense of loneliness and anguish, crying from inside to outside, soul to soul.” (Jones 2022). Will this sense of expressiveness satisfy the cluster property?

When discussing exceptions to each of the properties, Gaut states that “much of architecture and music is not concerned with the expression of emotion” (Gaut 2000, 33). It seems then that the work must be ‘concerned with’ the expression of emotion, not just have expressive qualities. It is unlikely that an AI will itself be concerned with the expression of emotion, as this is not a feature of current AI systems. However, that does not mean that a work made by an AI could not *seem* to express emotion, or to give viewers an idea of what it is like to experience an emotion, nor that an AI system could not be designed with emotion in mind.

At this stage of AI development, we are not seeing anything like emotional states being possible. However, there have been attempts to add an emotional component to machine images, such as the *emotionally aware Painting Fool* developed by Colton and colleagues (2008). We are also seeing so-called ‘empathetic algorithms’ which are designed to be sensitive to the emotional state of their users (Raamkumar and Yang 2022), and systems which researchers claim can identify emotional responses to images, particularly artworks, and express these in language (Achlioptas et al. 2021). Finally, there are AI-based art-making systems that claim to be able to create ‘emotional’ art, such as AIVA, an AI that

produces “emotional soundtrack music” (AIVA no date). Whether this system is successful in creating expressive music is debatable, but this certainly seems like it is creating works that are “concerned with the expression of emotion”. There is the added issue of where this concern comes from, and if it needs to come from the system rather than the designers of the system. Emotional or expressive considerations could plausibly be integrated into future AI art-making systems, particularly if this property is seen as key for art.<sup>2</sup>

So, if we want our AI to be expressing an emotion in making a work, this second property of art is not possible for current AI. This property may, however, be possible if we just want a viewer to find the work to be expressive, or if we allow ‘emotionally concerned’ works made by non-feeling AI to count under this property.

### **iii. being intellectually challenging (i.e., questioning received views and modes of thought)**

Works of art made by AI could be intellectually challenging to some extent. It may challenge us to wrap our heads around artworks being made by an AI, the images themselves may seem confusing, and we may try to glean information about the training data from the images we see. For example, the work of Google DeepDream can challenge us to guess on what the system was trained, understand how the system is recognising patterns, and consider what exactly this means for what the system ‘sees’ (Mordvintsev, Olah, and Tyka 2015). However, with AI works, much of the challenge will come from the fact that the work is made by an AI. As such, no autonomous AI system is likely to make works that are “questioning received views and modes of thought” without you needing to know that the work is an AI work, as that is a key factor in the work’s intellectual challenge. These AI systems do not have views or modes of thinking, and thus are not going to be able to question these through their work. Despite this, there is no requirement that the intellectual challenge of an artwork under the cluster account needs to stem from the work without context, or that the challenge must even be intentional. As such, this property does seem possible for an AI work.

### **iv. being formally complex and coherent**

There is no reason to think that AI could not produce formally complex works, particularly in the visual domain. Images from Generative Adversarial Networks (GANs), DALL-E and even older systems like DeepDream can produce colourful, multi-feature images. We have some evidence that AI can produce images that are formally complex, in comparison to some human-made artworks. Elgammal and colleagues (2017) conducted some initial interrogation into the responses to the images produced by their Creative Adversarial Network (CAN). I will draw on several of their findings during discussion of these properties. Elgammal et al. (2017) presented participants with images from the CAN versus abstract expressionist works and works displayed in *Art Basel 2016*.<sup>3</sup> In this study, participants were asked to rate the complexity of the images. Those images produced by AI systems were rated as higher on complexity than the human artworks in the study. The difference between the

---

<sup>2</sup> As I will argue later, the cluster account allows us to offer recommendations for the future development of AI for art.

<sup>3</sup> It should be noted that this study has limitations. The study was conducted online, most of the questions were asked to non-experts, and not all differences were investigated for significance.

sets was not large, but it may be sufficient for our purposes that the AI images were as complex as the human-made works (Elgammal et al. 2017, 16-17).

Formal *coherence* on the other hand may be more difficult for some AI. Though a system thoroughly trained on a narrow dataset will produce coherent, and seemingly representational,<sup>4</sup> images with frequency, a common way for AI systems to fail is to not produce a coherent image of a target depicted object (consider generative AI's problem with hands e.g. Hughes 2023). With increased diversity in the training set, or with a system designed to differ somewhat from the training set (such as AICAN), the images are not always clearly *of something*. Can they still be formally coherent?

The same study by Elgammal et al., discussed above, also asked participants to rate their agreement with the statement: "As I interact with this painting, I start to see a structure emerging." (2017, 17-18). Participants scored both the images produced by the GAN and CAN systems in the study higher than the human-art datasets. Seeing a structure emerging could point to formal coherence (in images); but the AI images being rated higher on this question than human-made images does not mean that they possess this property. Perhaps these art datasets are just not the kind to be "formally complex and coherent" (as is expected by the cluster account, some works of art will not have this property).

Hands notwithstanding, in general images produced by diffusion models (such as DALL-E and Stable Diffusion) seem to be more consistent in terms of producing convincing coherence and complexity. These images are generated in response to human text prompts and are based on a probabilistic model. In part because of this probabilistic basis, diffusion models are able to generate images which are structured in a coherent way. They do still make mistakes related to coherence, however (Romero 2022). For our purposes here though, it is sufficient that AI *could* produce coherent and complex works, and thus could produce works that meet this property under the cluster account.

## **v. having a capacity to convey complex meanings**

If conveying meaning involves communicating meaning from the artist to the viewer, it would appear that most art-making AI systems cannot do this. To do this, the AI itself would have to have a meaning in mind, which is not possible in the AI systems that we are examining, particularly generative algorithms. The more autonomous a system is (at least until we have systems capable of understanding) the less likely it is that the works have the capacity to convey *any* meaning, let alone complex meanings.

If we do not need the system to generate the meaning of a work itself, then it seems this property is already possible for some AI (depending on how complex the meaning is). For example, a diffusion-based AI like DALL-E can produce a relevant image when it is given a text prompt. In producing this image, the complex architecture of the system effectively takes the prompt, represents its meaning in a mathematical space, and uses this in the generation of an image. In this way then, the system is quite literally conveying meaning from one mode (text) to another (image). If successful, the image does indeed represent the meaning of the prompt (OpenAI 2022b). Whilst this kind of visual representation of meaning does not exhaustively capture what Gaut perhaps means by "capacity to convey

---

<sup>4</sup> I use the term 'representational' here simply to mean the images clearly depict a recognisable object.

complex meanings”, it could be a case of conveying meanings to a viewer. Given that the prompts for producing images in DALL-E are provided by humans, the system itself is certainly not conveying any meaning that it wants to. It *can* convey the meanings given to it by a human, but it cannot come up with those ideas for *what* to convey independently; it is reliant on a human user inputting text into the system. To use an example that is not a visual work, a written piece by an AI system such as GPT-3 could also include complex meanings that could be understood by a reader. However, this meaning is not connected to anything the AI system aims to convey; it is, in this case, produced predictively in response to a human prompt, much like elements of DALL-E.

AI works taken broadly could have a capacity to convey some complex meanings to an audience, particularly if they are prompt-based systems. If we require the meanings to come from the system itself though, this will not be possible for current AI.

## **vi. exhibiting an individual point of view**

Exhibiting an individual point of view may not seem very likely with an AI system that is producing artworks. First, it seems odd to say that they have a ‘point of view’, which we would typically use to indicate having a mind. To have a point of view to exhibit we would expect something to have a perspective on the world, even to have thoughts, feelings, and opinions on the world. All of these require a mind, and these AI systems do not have a mind. Second, AI systems are rarely thought of as individuals. This is partly due to their constant changes. Most AI systems employ machine learning (ML) algorithms. ML algorithms change constantly, with each round of learning resulting in alterations to the system. For example, in GANs and CAN, the process of learning alters weights in the system to improve future images. An image produced at an early stage of this process will be different to one produced later in the process. One might contend that humans too change, particularly through learning; artists will produce a variety of works over their careers, so how could we say they exhibit a single point of view? Despite this analogy, there is still a constancy that we cannot ignore about a human: they are one being, with a single physical manifestation. Furthermore, changes to humans occur slowly, and in response to environmental stimuli. Humans are embodied.

However, if we take ‘individual point of view’ to encompass producing images in a single style, AI may be able to achieve this. We can look at the work of a narrowly trained AI system, such as the GAN Obvious used to produce *Portrait of Edmond de Belamy* (the algorithm was trained and uploaded online by programmer/artist Robbie Barrat). This system produced multiple similar images, which look like a person with indefinite features seen through textured glass (Christie’s 2018).<sup>5</sup> Similar images were able to be reproduced using Barrat’s algorithm, which suggests that the algorithm developed and trained by Barrat had a particular style and subject it repeatedly reproduced in images. Unfortunately, though, this AI system has been trained to reproduce the kinds of portraits on which it has been trained. The fact that it produces multiple similar images is indicative of the narrow ‘experience’ that the AI has with images. It is not clear that this should count as a “point of view”. Furthermore, the selection of the training images is down to the human(s) training the system, in this case Barrat. This is how people exercise control over the images produced by

---

<sup>5</sup> It is not clear how much of the appearance of these images is due to selection, processing and printing by Obvious.

generative AI systems; a consistent aesthetic, and particularly a consistent subject-matter, is not evidence of the AI system having a point of view as much as the human having a specific aim for the system.

It is unlikely that an AI will be able to achieve this criterion, unless we consider “exhibiting an individual point of view” to be satisfied by “exhibiting consistency in style”.

### **vii. being an exercise of creative imagination (being original)**

Gaut offers two formulations of this property. The second, “being original”, should not be a barrier for AI systems producing works. AI systems can produce original works. This originality may be limited to the relevant domain; for example, a GAN is not going to produce anything that is not a digital image, though this system does produce novel images. It is worth noting that some generative AI systems may indeed have a problem with originality due to data over-fitting. This is where the system becomes too dependent on the training data, typically because there is not enough of it (Feng et al. 2022). When this happens in a GAN, it can result in the system reproducing images it has seen before. (Webster et al. 2019). A GAN that is not producing any novelty is a red flag for training issues. In the case of GANs, there are techniques to test for this replication and “Even when overfitting, most models will not reproduce perfect or trivially transformed copies of the training data” (Theis, Oord, and Bethge 2015, 6). Images produced by AI can also *seem* original to us. Again, we can turn to Elgammal et al. (2017). In one of their experiments, AI images were judged to be more original than the human artworks by participants. This originality is by design, particularly in the CAN wherein the system is designed to differ stylistically from training images.

*Imagination* on the other hand is more difficult for an AI system. I expect that we would anticipate having a mind (and mental representations) as central to imagination. If this is the case, AI systems are not going to be able to achieve this property, as the kinds of systems we have now (and for the near future) will not have minds. Despite the numerous concepts to which “imagination” seems to refer, Liao and Gendler state that broadly:

*To imagine* is to represent without aiming at things as they actually, presently, and subjectively are. One can use imagination to represent possibilities other than the actual, to represent times other than the present, and to represent perspectives other than one’s own. (Liao and Gendler 2020)

Generative algorithms can represent things:

In generative modeling, methods like variational autoencoders (VAEs) [KW13] and generative adversarial networks (GANs) [GPAM\*14] create latent spaces for modeling and synthesis of data [CN17]. Despite arising from different algorithms serving distinct purposes, these representations are all vector spaces of reduced dimensionality (relative to the input), intended to produce more general features that helpfully characterize the input. These representations are often referred to as latent spaces. (Liu et al. 2019)

Diffusion models have similar representational spaces. The latent space is constituted by “representations” of data on which the system has been trained. It does not contain all the

data but a compressed version of that data. It additionally represents relationships between the data that it has been trained on: “The concept of ‘latent space’ is important because its utility is at the core of ‘deep learning’ — learning the features of data and simplifying data representations for the purpose of finding patterns.” (Tiu 2020). This might be seen as akin to a form of imagination (though, perhaps not the creative kind). Rather, it might be more like a representation of something one has seen before (like picturing scenes from a film, and how they relate to each other). Given that this latent space is representative of the training data an AI is exposed to during learning, we might say that this is just representing things “as they actually, presently, and subjectively are” even though this kind of latent space is key to generating *new* images.

Despite this, some of the images produced by AI systems certainly seem fantastical. Take the examples made available by OpenAI on their DALL-E 2 Demo (OpenAI 2022a): if DALL-E 2 can produce an image of “Teddy bears shopping for groceries in the style of Ukiyo-e”, then should this not count as “not aiming at things as they actually are”? Whether we might consider these representations as (machine) imaginings will be dependent on how we want to characterise the pattern-finding in the latent space and, given the novel images that are generated from the latent space, if we are willing to call these “possibilities other than the actual”.

The focus of my response so far has been on “imagination”, but what of “creative”? In light of Gaut’s more recent work on creativity, the property of “being an exercise of creative imagination” might actually be akin to “being an exercise of creativity”. Gaut states in a later work on the link between imagination and creativity: “we have defended the existence of an a priori constitutive connection between imagination and creativity: imagination is suited of its nature to be the vehicle of active creativity.” (Gaut 2003, 289). Exercising creative imagination then could be equivalent to being creative in the cluster account. Whether an AI can be creative is a matter beyond the scope of this chapter.

### **viii. being the product of a high degree of skill**

If we are only to take skills as the kinds of things humans are skilled at (like wielding a paintbrush with precision, depicting an object accurately, or playing a piano well), then these AI systems are not skilled.<sup>6</sup> They may, however, become skilled in the sense of gaining the ability to do something well (e.g., generate images). We could characterise machine learning as a process of perfecting a skill for an AI system, as the AI is improving in its ability to produce images within certain parameters of success. In the case of a GAN system, the generator in the system generates images, which are judged by the discriminator. Feedback from the discriminator goes into the generative process, improving the images in the next round of generation.

### **ix. belonging to an established artistic form**

This is not too much of a stretch for AI. First, the majority of works considered here are digital images. Digital images are an established form of art, and they may also be emulating another established artistic form (paintings, photographs etc.). Second, AI is trained on

---

<sup>6</sup> Though some AI systems may be able to perform some tasks with greater speed and accuracy than humans (e.g., *AlphaGo*).



other artworks, and one of AI's clear limitations (in terms of creativity) is the extent to which it could ever go beyond the kind of works on which it has been trained. Systems such as GANs are designed to produce images that are a believable part of the set of images on which it has been trained, and predictive/probabilistic systems (such as DALL-E, GPT-n, AIVA etc.) are designed to continue on with what they have been prompted with in an expected way, also based on training data. Neither of these kinds of systems are going to produce works that are vastly different from the works on which they have been trained. This is not to say that an AI will not *ever* make works that do not belong to an established artistic form and depending how narrowly we take "artistic form" we may see AI systems produce different kinds of works (e.g., a new form of narrative work, or a new style of digital painting). However, it seems likely that most successful art-making AI will be producing works that belong to an established form. One might object that the fact that these are AI works in itself makes them another form of art. I am not convinced that this is a problem, as the focus in Gaut's property appears to be *form*, and not production, which would suggest a focus on the perceptible properties of the work.

#### **x. being the product of an intention to make a work of art**

A work 'being the product of an intention to make a work of art' of course requires intention. To have intention is often thought to require a capacity for mental states and with them a mind. If AI needs to have a mind in order to have intentions, then this will either not be possible at all (if we agree with Searle, 1980) or at least not possible in current AI systems. If we do not think that an AI needs to have a mind to have intentions, and instead substitute something like "aiming at a purpose" for intentionality, then some AI could achieve this criterion.

We could also adopt Dennett's (1989) proposal of the intentional stance, whereby the "truth" of the matter about a system's capacity does not matter. All that matters is the utility of adopting an assumption that the system is intentional. In order for this approach to be useful, we would need to know if people do indeed adopt an intentional stance to AI art. We do have some (albeit limited) evidence that people take an intentional stance to artworks made by some AI systems. Again, in Elgammal et al.'s (2017) study, participants were asked how much they agree with the statement: "As I interact with this painting, I start to see the artist's intentionality: it looks like it was composed very intentionally." (2017, 17). With this question, participants score the GAN and CAN images higher (more agreement with the statement) than the human artworks. There are limitations to this question. The clear issue for my purposes is that this is not asking whether the image appears to be intentionally a work of art. It is likely that this question is addressing formal qualities (I imagine a Jackson Pollock work would score low, yet it is clearly intended to be a work of art). However, if the works appear intentional in some way, this suggests that participants are taking an intentional stance towards these images.

Should apparent intentionality in a work of art count towards "being the product of an intention to make a work of art"? The work is the same, whether it is made by a human or an AI system. If it turns out that the work is made by an AI, does that mean that is not the product of an intention to make a work of art? If we take an intentional stance towards AI, then no; it would still be an intentional work of art. There is a clear objection here of course: many people do not sign up to a Dennettian view of intention.

Aside from this debate, AI art could utilise some intention derived from humans. There are two possible routes to this:

1. In the creation and training of the AI: for a domain-specific AI, a human has designed and trained the system for a purpose. AI systems that produce works akin to art will often be designed and trained with art in mind. In this case, the designer/engineer/artist is intending that whatever the AI produces will be art. The works produced by the system then are the product of the intention to make a work of art, even if there is an additional step in this causal chain.
2. In the selection of outputs from the AI: this is similar to found art or readymade artworks. In selecting the work made by an AI, a person may be, through that action, be intending to produce a work of art. The work would not count under this approach until this intentional action has been undertaken by a human, however.

If we are happy to utilise intentionality derived from humans, then this property is not a problem for works produced by AI. At higher levels of autonomy, however, we may be less willing to derive intentions from a human as in (1) (through the creation and training of the AI). If we are looking for an intention to produce a work of art in the AI system itself, this property is considerably more challenging, and the possibility of AI achieving this will be contingent on several factors, namely: the complexity of the AI, whether we think AI mind is possible, whether we are willing to accept a minimal account of intention, or whether we are willing to accept Dennett’s view.

Table 1: The cluster account of art – can AI works demonstrate each property?

|       | <b>Property</b>  | <b>Yes</b>     | <b>Maybe</b> | <b>No</b> |
|-------|--|----------------|--------------|-----------|
| i.    | possessing positive aesthetic properties   | ✓              |              |           |
| ii.   | being expressive of emotion  |                | ✓*           |           |
| iii.  | being intellectually challenging (i.e., questioning received views and modes of thought) | ✓              |              |           |
| iv.   | being formally complex and coherent  | ✓              |              |           |
| v.    | having a capacity to convey complex meanings   | ✓ <sup>†</sup> |              |           |
| vi.   | exhibiting an individual point of view   |                |              | ✓         |
| vii.  | being an exercise of creative imagination (being original)                               |                | ✓*           |           |
| viii. | being the product of a high degree of skill  | ✓              |              |           |
| ix.   | belonging to an established artistic form (music, painting, film, etc.)                  | ✓              |              |           |
| x.    | being the product of an intention to make a work of art                                  | ✓ <sup>†</sup> |              |           |

Note: these properties are taken from Gaut (2000, 28), and have been edited down into a list.

\*Dependent on how we take this property <sup>†</sup>Only likely in cases with human involvement

## ASSESSING THE CLUSTER PROPERTIES FOR AI

As I have shown, we can examine the list of properties Gaut provides us in the cluster account and assess whether AI can meet each in turn. By my assessment, several of these are possible, and more are possible with caveats (such as how we define elements of the property). Even more properties are possible as soon as we include humans in the equation (see table 1 for a summary). From this, it certainly appears that enough of the properties could be filled to count AI works as art (providing the necessary criteria is fulfilled – more on this below). One property which is particularly tricky for AI art is “exhibiting an individual point of view”. This property gets to the heart of the issue for AI. AI produces images that are, to a large extent, drawn from training on many other images. This is particularly true for diffusion models. It is then arguably the antithesis of something with an individual point of view. However, even on a more conservative view of the properties – let’s say AI works are lacking four properties – there is surely still scope for AI works to be considered art.

Gaut does not provide us with any constraint on the number of properties that he thinks sufficient for a work to be art under the cluster account. However we can use an example to demonstrate that a work can be lacking several and still be art. Take a work of art such as Yves Klein’s *IKB 79* (1959). This work is a vast, blue painted canvas. When measuring this work up against the cluster properties, we see that it may lack several. This work is not formally complex – it has little in the way of form at all. It’s debatable how much skill this work required to make; it is certainly less skilful than the ceiling of the Sistine Chapel, for example. The work is also debatably not particularly original in and of itself (Klein made nearly 200 blue, monochrome canvases in his lifetime (Howarth 2000)). It’s also debatable how much a minimalist work such as this could *exhibit* a point of view, even though the artist certainly had a point of view to their works. It’s not clearly expressive of an emotion; it is suggested that commercialism was at least a partial driver in the creation of these works (Howarth 2000). Whilst these points might be debated, let us take for a moment that Klein’s work did lack all five of these criteria. Despite this, would we deny it is a work of art? No – Klein’s works are art, and influential art at that. Thus, we could imagine that even lacking five properties in the list is not sufficient to deny a work is art.

I hope to have shown through my assessment of these properties how the cluster account of art might be beneficial for those interested in the possibility of AI Art: the cluster account provides a list of properties that we can discuss, debate the possibility of, and potentially consider for operationalisation. By this I mean that we can take a property and consider whether 1) current AI can do it, 2) future AI could do it, and 3) how we might make this possible for an AI. For example, we could prioritise formal coherence in future AI systems. To do this, we might want to focus on developing links between a structured visual system and image generation in our AI systems. The cluster account can give us more to work with than other accounts of art in developing future AI art-making systems.

## ADDITIONAL PROPERTIES

The discussion thus far has focussed on the properties of the cluster account and measuring AI against these. There is a fundamental assumption underlying this, however. This is that when AI makes art, it will be like human art, so much so that it will be recognisable by the same metric that we apply to human art. There is a reason for this. The motivation for this investigation is uncertainty about whether AI could produce art, and considerable scepticism

about the possibility that it ever could. There is an intuition for some, and more than an intuition for others, that art is a *human* endeavour. This framing is anthropocentric and ignores the possibility (maybe even a slim one) that whatever AI does will be quite different from what humans do. This is believed to be increasingly likely as we edge towards autonomy and beyond into human-level and beyond-human-level intelligence in AI (see e.g., Bostrom 2014). AI may well have different values to us, including in the realm of the arts, and these values may or may not end up appealing to humans.

There is a benefit here to a cluster account: it can be adapted. Gaut states that individual criteria can be disputed (Gaut 2000, 29). Furthermore, if an example could be found of a missing property that should be on the list, or a counterexample of a case that seems to be art but does not meet any of the criteria, then a new or missing property could be added:

There is no evident way that an object lacking all of the criteria could be a work of art; and even if a plausible counterexample could be produced, the friend of the cluster account could respond by *adding whatever seems like the relevant criterion to the cluster* — that is, she can respond by modifying the intent of the account, rather than its form. (Gaut 2000, 32-33) my emphasis)

Taken together, this suggests that we could make a case to add properties to this list where we have a counterexample case that *seems* to be art but does not meet enough of the properties in the list, and yet clearly has other properties that might be added to the list (without altering the form of the cluster account itself). In the case of an AI artwork that does not meet all, or even most, of the properties in the cluster account, there is a potential for us to adapt the account to alter existing properties or propose *new* properties that are unearthed by AI works.

A foreseeable case could be, for example, to add “information” to the property of “having a capacity to convey complex meanings”. “Having a capacity to convey complex meanings or impart complex information” could then include AI works’ ability to convey probabilistic information about artworks (as we see with DALL-E) or the ability to convey aggregate information about traits of a training set of images (as with GANs). A DALL-E image can tell us something about a large dataset that other forms of data visualisation or mere statistical analysis might not. For example, from discussion with artists using DALL-E 2,<sup>7</sup> DALL-E 2 struggles to depict a man wearing a dress and will opt for depicting a woman in a dress, or a man next to a dress. We could report on the percentage of images in the training data set that depict men in dresses, or the ratio of men to women depicted in dresses in the training data set, but this does not compare to the information we could gain from just looking at DALL-E images. Seeing a system that can depict “An astronaut riding a horse in the style of Andy Warhol” be unable to depict *a man in a dress*, reveals bias in the data, and the impact of that bias in one fell swoop. You do not have to agree that this is something that should count as a property of art (and perhaps by virtue of it being by an AI, you will not), but we can see in this example how an AI artwork might offer something different to human art.

---

<sup>7</sup> Artists working at Realdreams, an AI art collective operating in London (Realdreams 2022).

These AI systems are just today's technology, but already we can start to see how these "art"-making AI can provide something for art that Gaut's initial (anthropocentric) list of properties might not currently include. This is not to say that just *any* properties of AI works can be added to the account, and not just any alterations to the account can be made (Gaut 2000, 32). We would need to make a good case for the inclusion of a new property. Despite this, the cluster account gives those interested in the future of works made by AI a route to allow AI to contribute to what we see as art, not just play catch-up to human-made art.<sup>8</sup>

### THE NECESSARY CONDITION

My discussion of the cluster account thus far has ignored a glaring issue for AI: the action condition. Gaut's modified cluster account includes a single necessary condition, and that condition is that the artwork "is the product of an action". Can AI perform actions? Whether an AI can perform an action will depend on how we characterise actions and, in particular, what relationship we require between action and intention. This is not a settled debate in philosophy of action (Wilson and Shpall 2016). If the "action" required to make art a "work" does need to be intentional, then we once again reach an impasse for AI works. To do something intentionally, we would expect to need a mind, as intention is considered to be a mental state by many (Jacob 2019), though, as we have seen above, we do have some possible routes for AI intention.

Is the "action" in the cluster account to be an intentional one? This is not clear. The action does not need to include the intention to make a work of art, as that is listed separately as a disjunctive condition. Gaut reaffirms that this is not necessary in a later paper:

Those who claim that the disputed cases are not art may do so because they insist on, as a necessary condition for art, some feature, such as being the product of an intention to make art. But it is a mistake to insist on this as a necessary condition, as can be shown by considering less contentious cases of art, which lack this feature. (Gaut 2005, 281).

Gaut also makes it clear that the cluster account can accommodate various hard cases in aesthetics, such as Duchamp's readymades (Gaut 2000, 32, 35) and found objects (Gaut 2000, 29). Intention does not need to be present in the creation of the object, just in the selection of the work. In his discussion of Weitz's example of a found object, Gaut states:

A piece of driftwood in nature cannot express despair, nor can it be about anything (since it lacks even derived intentionality), but when selected for display in a gallery it can express desuetude and be about failure and decay (Gaut 2000, 29)

By selecting the object, we imbue in it the capacity for the other properties. This suggests that there is an easy way to avoid the issue of whether AI can act. In the case of much of today's AI art, a human will be selecting the outputs to share with the world, curating the AI outputs for wider dissemination (through online means, in galleries, in auction houses etc).

---

<sup>8</sup> I should note that I do not make an argument here that the cluster account of art is correct. I argue that it is the most useful for the purposes of assessing AI art.

If we do not want to worry further about whether AI works are art or not, we can again consider AI art to be much like found art, or readymades: it was not art before, but now it has been selected by someone. In selecting the image, the person has done the requisite (intentional) action to allow this image to count as a work, and with that has opened up the capacity of the image to hold more art properties (like the driftwood above). Similarly, we might think that an AI designed with the purpose of producing art would derive some intentionality from its own development, though perhaps with greater distance from intention to product than would be typical.

While this might offer an easy solution for us in the cluster account, it may be possible for AI to produce art without this solution. As I have shown above, an AI work could have several of the properties of the cluster account, whether it is properly thought of as a product of an action or not. If the action criterion is in place merely to open up the capacities of the object to have the other properties, then for AI this may not be necessary. It is already artefact-like, it is already the product of an action-like event, and it already has the capacity to have several of the disjunctive properties.

The inclusion of the action criterion is meant to establish that in order for something to be an artwork it must be a *work*. Here, again, is what Gaut says of the action criterion:

An artwork is the product of an action, preeminently of a making (an artifact), or a performing (a performance). It is *artworks* that are involved here, since something is in each case done. Hence being the product of an action is the genus of the artwork and is thus a necessary condition for something's being art. (Gaut 2000, 29).

The purpose of including “action” as a necessary criterion in the cluster account ensures that something is “done”. This then results in a product, i.e., the potential artwork. This seems to fit with the case of the products of AI. In adding an action condition, Gaut in effect *weakens* the standard account of artifactuality, rather than strengthens it. On the standard definition of an artefact as laid out by Preston (2020), artefacts must meet three conditions: 1) they must be produced with intention, 2) they must involve modified materials, and 3) they must be made for a purpose. Gaut, however, does not include any of these conditions. Besides a desire to include performances (the artefactual status of which has been hotly debated, see Preston 2020) Gaut also wishes to include objects that we might think were not produced with intention, or made for a purpose (e.g. found objects). AI does produce outputs, and (e.g.) make images. Can we not call these “works”? In the case of these outputs produced by AI, it seems very restrictive to call an AI product not art merely because the action which produced it might not be intentional enough, or the distance between intention and the product is too great (such as might be the case in an autonomous AI system).

In the case of AI creating art with human involvement (in the design of the system, or in the selection of outputs) the action criterion should be sufficiently fulfilled. In the case of future AI which autonomously chooses to produce art, it is not clear that we would deny their outputs status as works under the cluster account. Action is required to open up the possibility of achieving the other properties of the account, and many of these properties would be achievable by AI outputs.

## IS AI ART NOT ART?

In non-Wittgensteinian approaches to the definition of art (see e.g, Davies 1991 for an overview), if an AI art cannot meet the requirements of the definition, then it is not art. This is a limit of the typical definitional approach; the concept either does or does not apply. This is not quite the case for a cluster approach (even with the addition of the necessary condition). We have seen already that, according to Gaut, the cluster account includes some built-in flexibility for the possible future counterexamples, that is, we can modify the account (Gaut 2000, 33). We have also seen that the list of properties provided in the cluster account allows us to assess AI creations on multiple dimensions. There are further benefits to the cluster account, even if it turns out that AI art cannot meet the necessary criteria Gaut lays out. The cluster account is notable for how it fits with our intuitions not just about what *is* art, but also about what *isn't*.

Gaut discusses the adequacy of the cluster account by considering indeterminate cases, where there is considerable disagreement of whether a work falls under the concept of art:

if there are some objects to which the application of the concept is genuinely, irresolubly, indeterminate, then the account should reflect this too, rather than simply stipulating that the concept applies, or stipulating that it does not (Gaut 2000, 30)

Unlike definitional accounts of art, the cluster account does not deal with hard cases (AI art might be such a case) by mere categorisation. Instead, it offers us an explanation of the irresolution. For example, if AI works are not “art” (i.e., the concept as outlined by Gaut does not fully apply, and we have not modified it for the inclusion of AI works) then it may be explained as a borderline case:

the cluster account explains why some activities (such as cookery) seem to lie somewhere near the borders of art without clearly being art, since they share several properties of art (being the exercise of individual creativity, having a capacity to give sensuous pleasure), while also lacking other relevant criteria (since they have difficulty in expressing emotion and conveying complex meanings, and are not generally the product of an artistic intention). It is a signal advantage of the cluster account over the more straightforward definitions of art that it can preserve the hardness of such cases and allow us to explain what it is that makes them hard; such cases can be shown to be genuinely borderline and indeterminate. (Gaut 2000, 36)

Perhaps AI works are like cookery (or any other art-like case). They may be close to art, without clearly being art. Even if AI art does not meet the necessary criterion for being art (and one rejects my argument here), an example of an AI work might still meet several of the disjunctive properties Gaut lists, and thus we can explain why we might disagree with each other about our judgment of whether this example is art or not. This is better for us than AI art straightforwardly being rejected as art (and left there, as is the case with the other accounts). We can at least describe works made by (autonomous) AI as a difficult case, which lies close to the boundaries of art, and we can more easily justify discussion of the aesthetics of these works, why they seem artistic, or why some of us might still wish to label them as art.

The cluster account of art then can help us in several ways:

1. It can provide a framework for assessing what AI can or cannot currently do in terms of producing art, and this could help in the development of future AI.
2. It can potentially be adapted to accommodate what AI *can* do, through alterations or additions to the properties.
3. Even if AI cannot meet the necessary condition of the cluster account, the account can still explain why it seems like AI works are art, and why we might disagree about its status. AI art can be explained as a border case.

To conclude, in this chapter I have argued that a Wittgensteinian approach to art (through the cluster account) can have particular utility for AI art. I have argued here that (some) AI works could be art under the cluster account of art: they can potentially meet many of the listed properties featured in the account. With the intervention of a human, AI works can easily fulfil the necessary criterion. In the case of autonomous AI, AI works fulfil the apparent function of the action criterion. However, even with the rejection of the latter points, I have argued that the cluster account of art still has plenty to offer those sympathetic to the project of AI art.



## References

- Achlioptas, Panos, Maks Ovsjanikov, Kilichbek Haydarov, Mohamed Elhoseiny, and Leonidas Guibas. 2021. "Artemis: Affective language for visual art." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 11569-11579. <https://doi.org/https://doi.org/10.48550/arXiv.2101.07396>.
- AIVA. no date. "AIVA." AIVA Technologies. <https://www.aiva.ai/>
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Christie's. 2018. "Is artificial intelligence set to become art's next medium?". Last Modified 12 December 2018. Accessed 1 October 2022. <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>.
- Colton, Simon. 2008. "Creativity versus the perception of creativity in computational systems." *Proceedings of the AAI Spring Symposium on Creative Systems*: 14-20. <http://www.aaai.org/Papers/Symposia/Spring/2008/SS-08-03/SS08-03-003.pdf>.
- Davies, Stephen. 1991. *Definitions of Art*. Ithaca, NY: Cornell University Press.
- Dennett, Daniel C. 1989. *The Intentional Stance*. Cambridge, MA: MIT Press.
- Elgammal, Ahmed, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. 2017. "CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms." *arXiv*. <https://doi.org/https://doi.org/10.48550/arXiv.1706.07068>
- Feng, Qianli, Chenqi Guo, Fabian Benitez-Quiroz, and Aleix Martinez. 2022. "When do GANs replicate? On the choice of dataset size." *Proceedings of the IEEE/CVF International Conference on Computer Vision*: 6701-6710. <https://doi.org/https://doi.org/10.48550/arXiv.2202.11765>.
- Gaut, Berys. 2000. "The cluster account of art." In *Theories of Art Today*, edited by Noël Carroll. Madison, WI: The University of Wisconsin Press.
- . 2003. "Creativity and imagination." In *The Creation of Art*, edited by Berys Gaut and Paisley Livingston, 148-173.
- . 2005. "The cluster account of art defended." *The British Journal of Aesthetics* 45 (3): 273-288. <https://doi.org/doi.org/10.1093/aesthj/ayi032>.
- Howarth, Sophie. 2000. "Yves Klein IKB 79." Tate. Accessed 10 January. <https://www.tate.org.uk/art/artworks/klein-ikb-79-t01513>.
- Hughes, Alex. 2023. "Why AI-generated hands are the stuff of nightmares, explained by a scientist." BBC Science Focus. BBC. Accessed 10 January. <https://www.sciencefocus.com/future-technology/why-ai-generated-hands-are-the-stuff-of-nightmares-explained-by-a-scientist>.

- Jacob, Pierre. 2019. Intentionality. In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.
- Jones, Jonathan. 2022. "Incoherent, creepy and gorgeous: we asked six leading artists to make work using AI – and here are the results." *The Guardian*, 1 December 2022, 2022. Accessed 15 December 2022.  
<https://www.theguardian.com/artanddesign/2022/dec/01/six-leading-british-artists-making-art-with-ai>.
- Kennick, William E. 1958. "Does traditional aesthetics rest on a mistake?" *Mind* 67 (267): 317-334. <https://www.jstor.org/stable/2251530>
- Klein, Yves. 1959. IKB 79.
- Liao, Shen-Yi, and Tamar Gendler. 2020. Imagination. In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.
- Liu, Yang, Eunice Jun, Qisheng Li, and Jeffrey Heer. 2019. "Latent space cartography: Visual analysis of vector space embeddings." *Computer Graphics Forum* 38 (3): 67-78.  
<https://par.nsf.gov/servlets/purl/10172008>.
- Mordvintsev, Alexander, Christopher Olah, and Mike Tyka. 2015. "Inceptionism: Going deeper into neural networks." *Google Research* (blog).  
<https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>
- OpenAI. 2022a. "DALL·E 2." *OpenAI* (blog). <https://openai.com/dall-e-2/>.
- . 2022b. "DALL·E Now Available Without Waitlist." *OpenAI* (blog). 29 November 2022.  
<https://openai.com/blog/dall-e-now-available-without-waitlist/>
- Preston, Beth. 2020. Artifact. In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.
- Raamkumar, Aravind Sesagiri, and Yinping Yang. 2022. "Empathetic conversational Systems: A review of current advances, gaps, and opportunities." *IEEE Transactions on Affective Computing*. <https://doi.org/>  
<https://doi.org/10.48550/arXiv.2206.05017>.
- Realdreams. 2022. "Realdreams." Accessed 16 December 2022. <https://www.realdreams.io/>
- Robinson, Jenefer. 2005. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Oxford University Press.
- Romero, Alberto. 2022. "DALL·E 2, Explained: The Promise and Limitations of a Revolutionary AI." *Towards Data Science* (blog). 16 June 2022.  
<https://towardsdatascience.com/dall-e-2-explained-the-promise-and-limitations-of-a-revolutionary-ai-3faf691be220>.
- Roose, Kevin. 2022. "An A.I.-generated picture won an art prize. Artists aren't happy." *New York Times*, 2 December 2022, 2022. Accessed 12 December 2022.

<https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>

Searle, John Rogers. 1958. "Proper Names." *Mind* 67 (266): 166-173.  
<https://www.jstor.org/stable/2251108>.

---. 1980. "Minds, brains, and programs." *Behavioral and Brain Sciences* 3 (3): 417-424.  
doi:10.1017/S0140525X00005756.

Theis, Lucas, Aäron van den Oord, and Matthias Bethge. 2015. "A note on the evaluation of generative models." *arXiv*. <https://doi.org/>  
<https://doi.org/10.48550/arXiv.1511.01844>.

Tiu, Ekin. 2020. "Understanding latent space in machine learning." *Towards Data Science* (blog). 4 February 2020. <https://towardsdatascience.com/understanding-latent-space-in-machine-learning-de5a7c687d8d>.

Webster, Ryan, Julien Rabin, Loic Simon, and Frederic Jurie. 2019. "Detecting overfitting of deep generative networks via latent recovery." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*: 11273-11282.  
<https://doi.org/10.48550/arXiv.1901.03396>.

Weitz, Morris. 1956. "The Role of Theory in Aesthetics." *Journal of Aesthetics and Art Criticism* 15: 27-35.  
<http://www2.hawaii.edu/~freeman/courses/phil330/24.%20The%20Role%20of%20Theory%20in%20Aesthetics.pdf>.

Wilson, George, and Samuel Shpall. 2016. Action. In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.

Ziff, Paul. 1953. "The task of defining a work of art." *The Philosophical Review* 62 (1): 58-78.