Can AI Create with flair? Artistic Creativity and Artificial Intelligence

Can Artificial Intelligence (AI) be creative? In order to address this question, we need to understand creativity before moving to apply our chosen definition to the assessment of AI. With this in mind, this paper focusses on one component of one definition of creativity: flair. Flair is Berys Gaut's addition to value and originality, the components typically included in accounts of creativity. Flair is the basis of Gaut's rejection of AI creativity; but is he right to reject this possibility on these grounds? In this paper I argue that flair is possible for AI. I begin by outlining the flair condition. I then turn to assess whether flair is possible for AI, arguing that it is potentially achievable. I will then consider Gaut's objection to the possibility of AI creativity through flair. I argue if we are happy to accept that skill is sufficient for flair in place of understanding then flair is possible for AI. However, there remain two issues, which I will then turn to examine. The first issue is to clarify the relationship between skill and understanding, in order to establish whether they can indeed be pulled apart in flair. The second issue is to consider whether flair is inextricably linked to agency, thus is not possible for AI. I conclude that flair is possible for AI systems to achieve if we do not require understanding, suggesting that we can instead rely upon skill to help delineate creative processes in non-agents.

Introduction

Can Artificial Intelligence (AI) be creative? This question has been the subject of increasing public discussion, particularly with the proliferation of generative AI models used for artistic ends (Baxter 2024). Fortunately, this question is one that philosophical aesthetics has the tools to address. Aestheticians have put forward accounts of creativity, which we can use to assess the possibility of creative machines. This paper focuses on Berys Gaut's agential account. In particular, it examines Gaut's unique addition to his account of creativity: flair.

Why examine flair in the context of AI? Whilst agency has received attention in the discussion of creativity in AI (see Paul and Stokes 2024; Brainard 2025; Moruzzi 2025; Author forthcoming), flair has been relatively under-discussed.[1] This is despite the key role flair plays in Gaut's agential account of creativity. As Currie and Turner have described, Gaut's notion of flair

[1] For some brief discussion of flair in the context of AI, highlighting the anthropocentric nature of the concept, see Moruzzi (2021).

is "One of the most developed versions of creative agency" (2023, 119). Flair is a complex and somewhat idiosyncratic component of creativity, which may account for this relative lack of attention; yet its formulation offers direct objections to the possibility of machine creativity. Given this, and the prominence of Gaut's agential account of creativity (particularly in artistic domain), it is worth a more extended investigation in the context of AI.

This paper offers such an investigation. I will examine the flair condition of creativity, before turning to assess whether flair is possible for AI. I will directly respond to Gaut's rejection of the possibility of AI creativity through the 'understanding' component of flair, arguing that we should doubt the finality of this objection. I will propose that 'skill' can replace understanding in Gaut's account, particularly as Gaut at times suggests a disjunction between skill and understanding in his account of flair. I will argue that we have reason to think that skill does not imply understanding, and thus the two concepts can be separated. I will then go on to consider whether it is possible to have an account of flair which is non-agential.

Flair

For Gaut, creativity is a particular exercise of agency, consisting in three conditions (rather than the typical two-part definition): value, originality, and "flair". We might assume flair relates to something like "style" or uniqueness, or perhaps an aptitude. However, Gaut defines flair through the exclusion of four different cases that he argues should not be included in creative processes. First, he suggests that works made by accident cannot be creative. Consider an artist working in a studio who knocks over a pot of paint, which lands on a canvas laying on the floor. Even if this makes a beautiful and original artwork, Gaut argues that it cannot count as creative due to the work being a product of pure luck (2010: 1040; 2003). Gaut does not insist that *no* luck can be involved in a creative process; he thinks that capitalising on luck is permissible. As Gaut states, "serendipity is the skilful use of chance, not pure luck." (2010: 1040) Gaut also makes no mention of intention here and characterizes creative action without pure luck as "exhibiting at least a relevant purpose" (2010: 1040). Gaut does not add any requirement for intention here (if there is any need for intentionality is included in the requirement of agency).

Second, Gaut argues that following a mechanical search process cannot count as creative either. Gaut considers cases where "someone who produces something original and valuable simply by mechanically searching through all the possible combinations available to him" (Gaut 2010: 1040). He points to Novitz's discussion of Charles Goodyear's discovery of vulcanisation, which involved Goodyear mixing raw rubber with a series of available materials (Novitz 1999:

75). Gaut states that this example should not be counted as creative because it shows no skill or understanding in its application (2010: 1040).

This case is perhaps not the best example of mechanical search. According to Novitz the actual discovery of vulcanisation came when Goodyear *accidentally* dropped one of his mixtures onto a hot surface (Novitz 1999: 75), suggesting that this may be a case of 'pure luck' which is already excluded above. While Goodyear was indeed searching through some options, he was doing so fairly randomly, without any system in place. This seems more like 'blind variation', i.e., trying one thing after another as they are made available, and testing the outcome (selectively retaining) anything successful (see Helliwell 2021). We could ask whether a true mechanical search, systematically moving through a series of possibilities, would not involve some level of understanding? Gaut does phrase his case for exclusion as "mechanically searching through all the possible combinations available to him". He does indeed appear to have in mind a person systematically searching through all the possibilities one by one. This process, however, cannot produce results without *some* understanding. Either there will be understanding of the options for possible combination and the reasons behind testing them, or some understanding of the criterion of success (evaluative capacity is another criterion of flair, according to Gaut). It is also not clear then why Gaut thinks understanding or skill is the solution to excluding mechanical searches, as it is possible for a mechanical search to involve understanding (as opposed to a random search as Goodyear did, with an accident in the mix too, though even in this case Goodyear had in mind what he was looking for). This then is a somewhat confusing inclusion in the condition of flair, but Gaut proposes two further cases to exclude.

Gaut's third exclusion to creativity is a case of someone following exactly specified rules, such as a person following a paint-by-numbers in creating a painting (Gaut 2010: 1040). Deutsch (1991: 210-1), referenced by Gaut, argues that in completing a painting by numbers one is bringing into existence something that has already been *created* by another person. Gaut says, to avoid such a case counting as creative, there must be room for individual judgement for a process to be considered creative (Gaut2010: 1040). We should note here that only exact rule-following is excluded from creativity. Following rules non-exactly, or rules that do not specify everything to be done, should *not* be excluded from creativity. Creativity exhibited within set rules is still creativity, and at times constraining one's processes may even enhance creativity (Elster 1992; see also Levinson 2003 or Livingston 2009 for further discussion). In fact, examples of artists adhering to rules or constraints have frequently been examined in discussions of art and creativity. A central example would be *Litterature Potentielle*, where writers sought to give themselves strict constraints to follow in the creation of their writing (Symes 1999). Further

examples commonly discussed in the literature are the work of Igor Stravinsky, who also wrote of the need for constraint in music (Stravinsky 1997; Stokes 2008), and the forms and colours in the work of Mondrian (Stokes 2008: 226-8). It is not just philosophers and artists who have advocated for constraints as key to creativity, but also psychologists such as Patricia D Stokes (2005, 2008).[2]

It is not clear whether Gaut agrees with the possibility of constraint being essential to creativity. Gaut's view is compatible with rationalism about creativity (the idea that creativity is in some way rational, see Gaut 2012: 268) which would suggest that working within some rules is not incompatible with Gaut's position. Gaut does not argue against the possibility of rules having a place within a creative process (Gaut 2012: 260). This could lead us to conclude that whilst Gaut rejects the possibility of creativity occurring where *specific* rules are followed (such that no room is left for innovation), he is not arguing that *any* level of rule-following in a process precludes creativity from occurring. If we set our own rules to follow (much like some exercises of *Litterature Potentielle*) then one could surely be creative, perhaps not in following the rules, but in working creatively within the rules, or in setting the rules in the first place. This ultimately would show the 'degree of judgement' in applying the rules, as long as the creator determined the rule and its overall application.

Finally, Gaut wishes to exclude cases such as animals and young children who, when painting, will continue to add and add to the image unless it is taken away from them (Gaut 2010: 1040 citing Dutton 2009). In impressive cases of, for example, chimpanzees painting it is not the animal who ends the painting when it is done. Instead, a trainer removes the image at a point where the work is aesthetically appealing (and before the image is 'over done'). Gaut argues that in such cases the animal (or child) cannot have been creative since they lack the capacity to evaluate their own work and know when to stop creating.[3]

Gaut summarized the missing feature in the chimpanzee case as "an evaluative ability directed to the task at hand." (2010: 1040), however, I do not think this has to be evaluation in relation to finishing a work, or to evaluation of finished pieces only, as this would shift focus from a creative process to the product of this process. We could see evaluation in the process of

[2] It is worth noting that the view defended by Stokes is compatible with both Simonton and Boden's views, see Stokes (2008: 223).

[3] It does seem that there are some examples of chimpanzees who do (according to reports) demonstrate a desire to finish an unfinished work and refuse to continue when a work is finished (see Howell 2007, 374; Luty 2017, 383-385). This objection however might merely mean that some animals could be creative under Gaut's account (providing they can meet the other criteria).

making a work, not only in determining when it is complete. Consider a child who carefully painted a picture of a pink cat and then continued adding to the painting, eventually making a mess of wet paints. The process here, of producing an image of a fantastical animal, may nonetheless have been creative despite destroying the work after. I think therefore that Gaut intends to exclude cases where there has seemingly been no evaluation directed at the task at *any point* in the process.

Gaut puts proposes that flair will serve to exclude these cases. We can summarize Gaut's account of flair as requiring the following:

1) a relevant purpose (not accidental, or by pure luck)
2) some degree of understanding or skill (not merely using mechanical search procedures)
3) a degree of judgment (in how to apply a rule if a rule is involved)
4) an evaluative ability directed to the task at hand.

Flair and AI

Flair is, as highlighted by Moruzzi (2021) is an example of an anthropocentric conception of a feature of creativity. However, I will argue that it is indeed possible for AI to meet the requirements of flair. Let us look closer at the multiple aspects of flair and assess whether an AI could achieve these requirements.

Before examining each of the conditions of flair against an AI, it is worth noting that there are two potential ways to understand the condition of 'flair'. Whilst Gaut creates the positive condition of flair, featuring (in the 2010 version) four conditions that a process must meet to have been done with flair, it certainly appears that he created the condition of flair in order to *exclude* certain kinds of process (as opposed to including certain processes):

> the process of a thing's making involves, as I will say, *flair*. This should be understood to at least rule out the cases where I produce something by a mechanical search procedure, or in which I produce something purely by accident. (Gaut 2009: 86)

This is salient because in assessing whether an AI can meet each condition of flair, we can take each condition in two ways: whether an AI meets the positive condition (e.g., its work is produced with purpose) or whether it avoids the negative condition (e.g., its work was *not* created by accident). In some cases (such as in the case of mechanical search procedures) the positive and negative conditions may come apart.

1) a relevant purpose (not accidental, or by pure luck)

There is some randomness in systems like generative adversarial networks (GANs) and the creative adversarial network (CAN, see Elgammal et al. 2017) so we could consider this to be 'accidental' or 'luck'-based.[4] However, whilst generative machine learning systems such as GANs do involve random input, they are also based on learning. As the system is trained, it learns the distribution of the training data, and this impacts upon the output of the system. The random input built into the system will still have an impact on the final output. However, as soon as these systems have been trained on a data set, they are no longer producing images through *pure* luck, but through some chance and (as I argue below) some skill.

2) with understanding or skill (not by mechanical search)

Gaut wishes to exclude mechanical search procedures from creativity. This may present a problem for some kinds of AI system. Gaut discusses mechanical search in relation to AI in 'Creativity and Imagination' (2003: 277-278). Here Gaut discusses the role of imagination in creativity, and argues against its role as allowing people to search through possible options:

> For consider Deep Blue, the chess computer which beat Kasparov in 1997. Deep Blue really does survey vastly more possible positions than any human could, and selects from them the one most likely to win the game. Deep Blue has in this sense a powerful imagination. But the problem is that it is the epitome of an uncreative way to play chess: it mechanically searches through the possible positions to arrive at the best. Kasparov in contrast, plays chess creatively, but cannot do so by surveying the vast numbers of possibilities that Deep Blue does. Creativity is precisely not a matter of a powerful imagination, in the sense of an ability to search through vast numbers of possibilities. (Gaut 2003: 277)

Gaut clearly states then that IBM's Deep Blue plays chess without creativity. What of Deep Blue's spiritual descendant, the Go playing algorithm of DeepMind, AlphaGo (DeepMind n.d.)? AlphaGo is an AI system designed to be able to (learn to) play the game of Go, a game previously thought too challenging for a computer to be able to play (let alone beat a human champion). AlphaGo went on to win against humans, including a high-profile series of games against a world champion, as documented in the film *AlphaGo* (2017). In a game against 9-dan world champion Go player Lee Sedol, AlphaGo made a surprising move – move 37 – which would never be played by a human opponent (*AlphaGo* 2017). Many lauded this move as evidence of AlphaGo's creativity, including its opponent: "I thought AlphaGo was based on

[4] Whether this is indeed true randomness is debatable, however it is effectively randomness to the system itself.

probability calculation and that it was merely a machine. But when I saw this move, I changed my mind. Surely, AlphaGo is creative." (Lee Sedol, quoted by DeepMind, n.d.).

By Gaut's account, it is unclear whether something like AlphaGo's much lauded 'move 37' would count towards having flair. AlphaGo systematically searches through options by way of a decision tree, so it is in some ways conducting a mechanical search. The process by which it works is a Monte Carlo Search Tree, combined with policy networks and a value network. The policy network will suggest promising actions to take, and the value network will evaluate the board positions. This is coupled with the Monte Carlo Search Tree which simulates games to find the best action, i.e., the action with the highest probability of leading to a winning game. In other ways, however, AlphaGo is not *merely* a mechanical search procedure, despite a search through various possibilities being a key part of its process. The system has indeed searched for possibilities, but it has also simulated games and then evaluated the possible options before selecting the best possible course of action. It is also not possible for the system, advanced though it is, to model *all* possible games, so some pre-selection has occurred. I get the sense that Gaut *would* wish to exclude AlphaGo from creativity on the basis of this search procedure, but I am not convinced we can exclude it so easily. AlphaGo does not follow Goodyear's search process, as it is far more informed than Goodyear, and it does *not* search through a series of random options.

This said, I doubt that 'move 37' is the hill to die on for AI creativity. We can look to other AI systems that do not work via any kind of mechanical search. Generative Adversarial Networks (GANs), Creative Adversarial Networks (CANs) and, the latest technology on the AI art scene, Diffusion Networks (such as those used by DALL-E, Midjourney, and Stable Diffusion) do not rely on mechanical search procedures to generate images. If Gaut wishes to ensure that mechanical search is not included in creativity, then in these cases we need not be concerned. If, however, Gaut intends for flair to include understanding or skill *regardless* of mechanical search, then we have further work to do.

Gaut has written more on the relationship between creativity and skill (Gaut 2009). He wishes to include skill as part of flair, in both excluding mechanical search and in excluding cases of pure luck (Gaut 2009: 85-86). Gaut here offers us a partial account of what skills are:

> First, the capacity in some domain is special (not universally shared): we tend to talk of people as skilled in some activity when they have an ability that is not possessed by everyone who engages in that activity. Second, we talk of skills as a kind of accomplishment (the words 'skill' and 'accomplished' are near-synonyms). Third, it

> makes sense to talk about practicing one's skills: in practicing music making, I thereby practice my music-making skills. Finally, skills are something that one can learn, so we oppose skills to purely natural abilities. Breathing is not a skill, because it is not something one learns; but music-making is a skill, since it is learned, and (arguably) walking is also a skill, since that is something that is learned too. (2009: 95)

So, we could summarize this account of skill as follows:

i) The capacity is *special* (i.e. it is not universally shared).
ii) It is a kind of accomplishment.
iii) It can be practiced.
iv) It can be learnt.

We might wonder whether some AI systems, particularly those which employ deep learning have some level of skill. Do they meet Gaut's criteria?

i) Not universally shared: The ability to generate images (for example) is not something universally shared by algorithms. The ability to generate *good* images (either judged as images that could be part of the training set as in GANs and CANs, or as judged by humans) is also not universal. Only some systems, and (perhaps more pertinently) only some iterations of systems can produce images judged to be good.

ii) Accomplishment: We might not necessarily consider an AI system itself to be accomplished, but this could be about our unwillingness to praise a machine. However, the outputs of many AI systems are praised widely, as we have seen in the news, online, and in exhibitions and auctions. If we are able to attribute responsibility for the work to the AI (this will require some level of autonomy at least) then we may consider the AI to be accomplished.

iii) Practice: I will take practice to mean repeated exercise of an ability. Many AI systems do indeed repeatedly exercise their abilities, either continually generating (and, for GANs and CANs, assessing and improving upon) their images or (in the case of DALL-E) generating multiple images in response to prompts. We might also require that the repetition results in improvement. In the case of adversarial networks, these repeated exercises will result in improvements.

iv) Something that one learns: it seems easy to take at face value here that machine learning involves learning.[5] One could object that a deep learning algorithm like GANs or diffusion models do not learn to make images as they are programmed to make images; thus, they do not learn the skill of 'image-making'. We could respond, however, that these systems learn to make images of a certain kind or quality.

I have argued that an AI system could meet each of the criteria of skill. In some places though, it seems that it is not understanding *or* skill which Gaut includes in 'flair', but just understanding. Here, it seems that either understanding *or* skill will satisfy Gaut: "Goodyear's discovery does not count as creative since it displays no understanding *or skill*" (Gaut 2010: 1040, my emphasis). Elsewhere, it seems to be *understanding* that is emphasized and not skill: "in short, the kinds of actions that are creative are ones that exhibit … some degree of *understanding*." (2010: 1040, my emphasis). Whether an AI system can have *understanding* will need further interrogation, as this is a far more philosophically-laden issue. The potential lack of understanding in AI systems is discussed by Gaut in relation to creativity. I will examine this further below, as Gaut's brief discussion of AI creativity can be taken as an objection to the argument I offer here. Let's return to assessing AI against the components of flair.

3) a degree of judgment (in how to apply a rule if a rule is involved) (not specific rule-following).

If we are to have some level of autonomy in our AI system, this should ensure that the system is not *merely* following rules. Not following specific rules, for an AI, could be interpreted as not only following pre-determined rules, i.e., not only following rules assigned to it from an external source. Some might object that as an AI system made up of algorithms, is it comprised solely of rules and thus is, in its very nature, unable to exercise judgement in applying rules. However, it's not clear that current AI can be simplified as following rules in this way. There is a key distinction to be drawn between "GOFAI" (Good Old-Fashioned AI), and contemporary AI systems which are made up of deep neural networks. GOFAI systems are programmed with explicit rules in mind. In contrast, AI built on neural networks are *trained*. It is this training which determines the system's outputs, not pre-programmed rules. As Langland-Hassan (forthcoming) notes, this distinction is significant for considering whether AI systems can be said to be following rules:

[5] Indeed, it is described as such widely, and modelled on human learning (Michalski, Carbonell and Mitchell 1983) though the idea that machine learning is 'real' learning is not without debate (Bringsjord et al. 2018).

> The important point for our purposes is that training the model does not involve providing it with a preformulated set of rules to follow in response to each input it may receive. One can think of the network as discovering the rule (or rules) for categorizing inputs on its own, yet where the "rule" devised may not be something straightforwardly expressible in a natural language, but that is instead scattered across the discrete activation and connection levels of thousands, millions, or even billions of internal connections among nodes. (Langland-Hassan, forthcoming: 4)

We can therefore appeal to the separation between the designer of the system and the AI itself when considering whether AI is merely following rules. The non-representational nature of the nodes and their connections and the high numbers of both nodes and connections in complex deep learning algorithms leads to a 'black box' problem, whereby even an expert computer scientist might not be able to explain why the system produces the outputs it does with any detail. Even with access to the algorithm's code, the non-representational nature of the system means that we still may not know what is occurring within the system, as we cannot interpret the numbers that we can see (see IBM 2024). Developers therefore cannot predict the system's outputs. If we have a level of autonomy in the system, there will be behaviours that are not determined by a pre-existing set of rules. They may, however, be determined by the AI itself through deep-learning and self-evaluation.[6] In this case then, we could claim that judgement in application of any rule has been involved in the self-altering of the parameters of the system.

As noted above, we need not remove all rules, merely ensure that there is not only following of rules determined by others. This will not work for all systems; we will likely require some level of autonomy in the system. However, as long as the system was determining its own rules for creating works (and not merely those pre-determined by a human designer, for example) then we should satisfy this condition of flair.

4) an evaluative ability directed to the task at hand

In the case of some AI systems (notably GANs) evaluation is central to their architecture. GANs are generative deep learning algorithms made up of two neural networks: a generator and a discriminator. The discriminator is trained on a set of training images, and learns to distinguish images from the training set. The generator begins producing images, initially randomly (with input from a noise vector). These images are fed into the discriminator, which will judge whether the image could be part of the training set or not, given what it has learnt about the

[6] In the case of GANs and CANs, we also have the addition of randomness (or pseudo-randomness).

training data. The discriminator produces a score for each image it processes; if the discriminator determines that the image could not be from the training set, it will give the image a low score. If it determines that the image is plausibly part of the training set, then it will give it a high score. These scores are fed back into the generator, which uses them to adjust weights in the network. These weights impact the next round of images that are produced by the generator. Gradually, through the implementation of this feedback, the generator will produce images that are increasingly similar to the training set of images. (Goodfellow et al., 2014).

In the case of GANs, there is an explicit evaluative process involved in the process of generating images. The discriminator evaluates the outputs of the image generator in the system. These outputs are evaluated against success criteria and given a score. This score is then used to improve future iterations of image generation. Consequently, we could certainly say that there is an evaluative ability at play in (at least some) AI systems.

Given my assessment above, it seems that AI systems *may* be able to achieve flair. However, when writing in 2010, Gaut suggests that flair provides reason to reject the possibility of machine creativity. Let us turn to his objection next.

The Searlean Objection

In his critiques of Boden's account of creativity, Gaut addressed the possibility of creative computers. Gaut claims that his account of creativity is incompatible with AI. This is (perhaps surprisingly) *not* due to the agential aspect of his account.[7] It is instead due to the element of understanding, provided through flair:

> If, as I will argue shortly, creative activity requires some degree of understanding, then, if computers cannot exhibit understanding (Searle), they cannot be creative. And if not all domains of human understanding can be specified in terms of rule-governed operations (Dreyfus 285–305), then in these domains computational theories of creativity that are based on classical AI are not possible (Gaut 2010: 1037)

The second point here I will dismiss. First, there is no reason to argue (and Gaut does not seem to) that *all* domains of human understanding would be needed to provide sufficient understanding for AI to be creative. As Gaut says, in any domain where we cannot specify our

[7] See Author (forthcoming) for further discussion.

understanding in a way that can be processed by an AI: "*in these domains* computational theories of creativity… are not possible." (my emphasis). This implies, though, that any kind of understanding that *can* be specified in the relevant way can have creativity in the relevant domain.

Second, Dreyfus, here quoted by Gaut, was writing from a Heideggerian perspective between the early 1970s and early 1990s (dependent on edition). Dreyfus argued that knowledge is tacit and cannot be fully articulated and thus cannot be given to an AI (Fjelland 2020). As summarized by Russell and Norvig (2010):

> Essentially, this is the claim that human behavior is far too complex to be captured by any simple set of rules and that because computers can do no more than follow a set of rules, they cannot generate behavior as intelligent as that of humans. (Russell and Norvig 2010: 1024)

As Russell and Norvig state, this issue was indeed a serious one. However, since Dreyfus's writing, the design of AI systems has changed to incorporate solutions to this and other problems (2010: 1025). Of course, to be fair to Gaut, also writing in 2010, the latest advances of AI were not available for consideration (though the deep learning revolution was already underway, see Sejnowski 2018). If we set aside the conceptual baggage of Dreyfus's approach, we can examine the shift in the methods of AI design from rule-focused approaches to bottom-up approaches (such as deep learning algorithms). While it certainly is the case that we have not (yet) created a General Artificial Intelligence (AGI), we have made considerable progress in terms of machine learning techniques, computational capacity, and the tasks that AI can perform (Bryson 2018). Some domains of ability that were previously thought out of reach for AI are now within reach, including the generation of language and images. The basis of these advances in machine intelligence is a change in how they are designed, away from the classical AI architecture familiar to Dreyfus.

Let us return to the first of Gaut's claims above: "If, as I will argue shortly, creative activity requires some degree of understanding, then, if computers cannot exhibit understanding (Searle), they cannot be creative." (2010: 1037). Gaut seems to indicate an acceptance of Searle's position on understanding in computers. There are (at least) three possible responses to Gaut, which may still allow for AI flair, and thus the potential for AI creativity. The first would be to reject the necessity of understanding for flair, and thus creativity, entirely. To do this we would effectively abandon Gaut's account of creativity. We could take this option and instead adopt Boden's account of creativity (Boden 2004), the Darwinian account (e.g. Simonton 1999) or some other account of creativity. I will not take this option here as my focus is the agential

account. The second option will be to accept Gaut's characterisation of flair in terms of excluding certain kinds of cases from creativity but answer the need to exclude 'mechanical searches' with something other than understanding. I have indicated this approach above, with the replacement of understanding by skill (Gaut's own term). The third option will be to reject Searle. Understanding is key to Gaut's account of flair, but Searle's argument that AI cannot achieve understanding is not.

Rejecting Searle's argument, as many have done (e.g., Boden 1988; Harnad 1989; Sprevak 2007), is our third possible response to Gaut. If Searle is *wrong* about the impossibility of understanding for AI, then Gaut's suggestion that this is reason to think AI creativity may not be possible is also incorrect. We might, as a weaker position, argue that whilst Searle's argument fails, we do not yet know whether understanding is possible for AI (and thus we cannot yet conclude whether creativity is possible for AI either).

Let us now turn to Searle's argument, so that we might reject it. In Searle's 1980 paper 'Minds, brains and programs', he argues against the possibility of 'Strong AI'. Searle understands Strong AI as the idea that an AI that is programmed in the right way will have a mind (Searle 1980; 1999). This is in contrast to 'weak' AI, where a computer programme can potentially simulate mental processes but does not actually have mental states (Searle 1980; 1999). Searle makes his argument using the 'Chinese Room' thought experiment. Despite Searle's considerable legacy, many have rejected Searle's argument that AI cannot achieve understanding.[8]

Whilst Searle's original argument is now over 40 years old, its legacy remains. Some think there has been no conclusive response to Searle enough to put the Chinese Room to bed (Cole 2020). Many, however, have put forward criticisms and objections to Searle's argument. Some examples include the systems reply, the robot reply, the brain simulator reply, the other minds reply, and the intuition reply (Cole 2020). Boden offers a response to Searle, which Gaut does not mention even in using Searle to reject her reasoning about AI creativity (Gaut 2010). Boden (1988) follows the lines of criticism of the systems reply and the robot reply. The systems reply argues that while the man in the room may not understand Chinese, that is not the same as whether there is understanding in the system as a whole. This, it is argued, is more akin to how an artificial intelligence would work: as an entire system. The robot reply argues that with the addition of embodiment and perception the Chinese Room would not hold; the reason the man has no understanding is because he is locked in his room with the instruction book, not out in

[8] See Cole (2020) for a comprehensive list of philosophers and their various responses to Searle.

the world. This second response is less helpful for us unless we wish to develop embodied systems.

I will not explore the responses to Searle further here. I seek only to cast doubt on Gaut's acceptance of Searle, whose position on machine understanding is far from settled. If we can reject Searle, then we do not need to accept that understanding is out of reach for computational systems. If we cannot reject Searle, then we might still be satisfied to use skill in the place of understanding when considering AI. However, we need to clarify the relationship between skill and understanding to secure this as a solution.

Skill and Understanding

There is a potential issue for the use of skill in place of understanding. Is it not clear from the discussion so far what, if any, is the relationship between skill and understanding. It may be that that skill may *requires* understanding. If this is the case, then skill cannot be considered as an alternative to understanding in flair. It will always be skill *and* understanding, not ever skill *or* understanding (as skill will also imply a kind of understanding and thus face the same objection as above). Thus, an account of flair which includes skill in the place of understanding will not provide a response to the Searlean objection.

Unlike knowledge and action, skill has not played a central role in analytic philosophy (Stanley & Williamson 2017). Much debate about skill has been focussed on its relationship with knowledge, rather than explicitly with understanding. Perhaps the clearest suggestion of a connection between skill and understanding comes from the characterisation understanding as linked to a set of abilities, such as the ability to explain the object of our understanding (e.g. Hills 2009). Under this view, understanding is very similar to knowing-how (Gordon no date). Similarly, recent debates on knowledge-how have focussed on skills and their nature (Pavese 2022). There are though (at least) two routes to doubt that skill requires understanding. The first, is to point to the clear intellectualist position underlying such a connection. There is an ongoing debate in the philosophical literature regarding intellectualism about skill. This debate centres on the question of whether skill is always guided by propositional knowledge, and indeed there is pushback to this view. For example, Robertson and Hutto (2023) put forward the possibility of intentional and non-intentional skills, arguing that "Allowing that that non-intentional doings, simply and straightforwardly, can qualify as skilled, as we have suggested above, yields clear theoretical advantages." (2023: 143). Whilst their view that skill does not require knowledge (and instead can

be linked to habits) is not uncontroversial, if they are correct about the possibility of non-intentional, non-knowledge-based skills, the idea that skill requires understanding (seen as a higher level of epistemic requirement than knowledge) is unlikely.

What would a skill look like, if it does not require understanding? I think, in fact, this may be quite intuitive in a creative context. For example, do we need a performer to have understanding to be a great actor? Intuitively at least, it seems that one need not have propositional knowledge, or explanatory ability, in order to be a great performer. The adage that great sports people do not always make great coaches points to a similar idea. Skill in a particular domain does not always translate to the level of understanding we might expect in someone who can explain what they do and thus can train others. The reverse also seems true; a coach may have none of the skills needed to play a sport and yet have a great understanding of it.

A second point against thinking that skill requires understanding is that the reverse has been argued: understanding requires skill. This stems from the argument that understanding (as opposed to knowledge) has final value (Grimm 2024). In particular, Pritchard has argued that understanding is valuable because it is an epistemic achievement:

> By contrast, Pritchard believes that understanding always involves strong cognitive achievement, that is, an achievement that necessarily involves either a significant exercise of skill or the overcoming of a significant obstacle. (Gordon no date)

Under this view, understanding is a cognitive achievement, and achievement in this case involves skill. Whilst skill may be necessary to achieve understanding, it does not follow that one cannot have skill without understanding. It seems more plausible that skill is a precursor to understanding.[9]

The discussion of the links between achievement and skill, understanding, and creativity, requires a fuller examination, but this is for another paper. For my purposes here, I will hold the position that skill does not *require* understanding, and thus does not fall foul of the same Searlean objection that Gaut poses.

Flair and Agency

[9] If correct, this might lead to some concern from those worried that AI is heading towards the possibility of understanding.

I have put forward the case that AI can meet the requirement of flair (bar understanding, if skill is insufficient). However, flair is a key component of Gaut's *agential* account of creativity, and thus it stands to reason that flair and agency may be inextricably linked. For example, Paul and Stokes (2018), who also favour an agential account of creativity, articulate Gaut's account of flair as follows:

> Put most simply, the proposal says that the right kind of process for creative production is one that involves, in some non-trivial way, the agency of the creator. Gaut (2003; 2010; this volume) offers one articulation of such a proposal, as he maintains that the creative process involves "flair," which is shorthand for a number of agential features. (Paul & Stokes 2018)

Paul and Stokes suggest here that flair represents agential features. We could take this to imply, therefore, that flair forms part of the requirement for agency in creativity. However, given my discussion above, I would argue that each of these could be met by an AI system which is not agential (and thus is not following an agential process).[10]

I am not the only one to suggest that flair, or an analogue of it, can be achieved in non-agential processes (or by non-agents). Currie and Turner (2023) have recently argued that agency and flair can be riven. In their paper, they put forwards an analogue of flair without agency, accounting for the apparent flair exhibited in evolutionary creation. At the centre of their argument is the idea that flair can be understood as a vehicle of creativity, rather its source. As Currie and Turner point out, whilst flair may be a component of the process of creativity in humans (in terms of producing products that are original and valuable) this does not mean that flair cannot operate differently to this. They write:

> Understanding and goals are often crucial aspects of how original and valuable products are generated, but this is how flair operates in agents like us, just as the imaginative faculties are often used. In non-agential processes like evolution, original and valuable products are generated by sensitive, exploratory, combinatory and transformative cycles of mutation, birth and death, and so forth. (2023: 123)

Given this, Currie and Turner' (2023) argue that evolution exhibits *non-agential* flair:

> On the account suggested thus far, some process can be creative when it has the capacity to act with 'flair', where by flair we mean a process which (1) uses Boden's creative

[10] I have argued elsewhere that the agency argument against AI creativity has not been convincingly secured (see Author, forthcoming – redacted for blind review). See also Moruzzi (2022).

> processes, (2) is sensitive to context, (3) is spontaneous, (4) prompts inquiry. In agents like us, (1)–(4) are generated by goal-orientation and understanding, and thus deserve praise. But in other processes (we've highlighted evolution) 1–4 are generated without those agential features, so are not the proper targets of praise and blame. (2023: 123)

To follow Currie and Turner's line of reasoning, we can understand AI as also following a non-agential process with flair. Unlike Currie and Turner, I do not reject goal-orientation. Indeed, I have argued elsewhere (Helliwell, forthcoming) that AI cannot be agential (understood as requiring intentionality), but rather can 'act purposively', thus can be goal-directed.[11] However, like Currie and Turner, I would also reject the idea that understanding is necessary for flair. As Currie and Turner (2023) argue, there is little lost when we jettison praise and blame (which, they suggest, are connected to intuitions that understanding is necessary for creativity), opening the path to non-agential creativity (2023: 113).

As I have suggested above, we can still exclude 'uncreative' processes by utilising skill in place of understanding. Little then is lost, and we can still talk sensibly about creative processes. An AI process conducted with flair would not require understanding (as this is possible only for agents) but would be process that is not based on pure luck, is conducted with judgment and skill, and involves evaluation. What I am suggesting here is, I think, less radical than Currie and Turner's position. I am suggesting that AI could meet an account of flair in creating outputs which mirrors Gaut's; the key distinction will be that AI cannot do so with full agency. it will, similarly to Currie and Turner, be a non-agential form of flair, precisely because it is not done by an agent.[12]

Conclusion

The question of whether AI can be creative requires careful examination of accounts of creativity to answer. Here, I have examined one component of creativity. I have argued that Gaut's condition of flair *can* accommodate machines. I have examined Gaut's account of flair, including the positive and negative construal of each component of the condition. Through this, I argued

[11] I suggest that this purposiveness is sufficient to meet the responsibility requirements of creativity without necessitating intentional mental states.
[12] I am not suggesting here that AI is more akin to evolution in terms of its creative process (for an analysis of this, however, see Helliwell 2021).

that each component is potentially achievable for AI, aside from understanding. However, I suggest that skill can fulfil the same role as understanding in excluding 'uncreative' processes from this account. I have also examined a possible objection to AI creativity from Gaut himself, stemming from the condition of flair. Gaut suggests that AI creativity is not possible because AI cannot understand and thus cannot produce work with flair. Whilst I have suggested that skill can suffice for flair, I argue that Gaut's objection hinges on the acceptance of a highly disputed position, which is not necessary to accept within the agential account of creativity. I then go on to assess whether my proposal to keep skill and jettison understanding is tenable, or whrther the two concepts are inextricably linked, concluding that we can speak of skill without requiring understanding. Finally, I examine the centrality of agency in the characterisation of flair, arguing (following Currie & Turner 2023) that we can have a non-agential account of flair.

Of course, this argument alone is not sufficient to argue that AI creativity is possible. Other conditions for creativity (such as being an agent) are also argued to not be possible for AI. This may be the case. Here I have simply argued that flair is possible for AI systems to achieve if we do not require understanding, suggesting that we can instead rely upon skill to help delineate creative processes in non-agents.

References

AlphaGo (2017) Directed by Greg Kohs [Film]. Netflix.

Baxter, C (2024) 'AI art: The end of creativity or the start of a new movement?' *BBC Futures.* 21 October 2024. Available at: https://www.bbc.com/future/article/20241018-ai-art-the-end-of-creativity-or-a-new-movement (accessed 1 November 2024).

Boden, M. A. (1988) 'Escaping from the Chinese room', in Heil, J. (ed.), *Computer Models of Mind.* Cambridge: Cambridge University Press.

Boden, M. A. (2004) *The Creative Mind: Myths and Mechanisms,* 2nd edn. London: Routledge.

Brainard, L. (2025) 'The curious case of uncurious creation'. *Inquiry,* 68(4), pp.1133-1163.

Bringsjord, S., Govindarajulu, N.S., Banerjee, S., Hummel, J. (2018). 'Do Machine-Learning Machines Learn?' In Müller, V. (ed) *Philosophy and Theory of Artificial Intelligence 2017. PT-AI 2017*. Studies in Applied Philosophy, Epistemology and Rational Ethics, vol 44. Springer, Cham. https://doi.org/10.1007/978-3-319-96448-5_14

Bryson, J. J. (2019) 'The past decade and future of AI's impact on society', in *Towards a New Enlightenment? A Transcendent Decade*, 11, Madrid: Turner, pp. 150-185.

Cole, D. (2020) 'The Chinese room argument' In Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy.* Available at: https://plato.stanford.edu/archives/win2020/entries/chinese-room/ (Accessed: 1 November 2022).

Currie, A. and Turner, D. (2023) 'Creativity without agency: Evolutionary flair & aesthetic engagement.' *Ergo Open Access Journal of Philosophy,* 10. 110–134. https://doi.org/10.3998/ergo.4633

DeepMind (no date) *AlphaGo.* Available at: https://www.deepmind.com/research/highlighted-research/alphago (Accessed: 1 October 2022).

Dreyfus, H. (1992) *What Computers Still Can't Do: A Critique of Artificial Reason.* Cambridge, MA: MIT Press.

Dutton, D. (2009) *The Art Instinct: Beauty, Pleasure, and Human Evolution.* Oxford: OUP

Elgammal, A., Liu, B., Elhoseiny, M. and Mazzone, M. (2017) 'CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms.' *arXiv.* Available at: https://doi.org/10.48550/arXiv.1706.07068

Elster, J. (1992) 'Conventions, Creativity and Originality', in Hjort, M. (ed.) *Rules and Conventions: Literature, Philosophy, Social Theory,* Baltimore, MD: Johns Hopkins University Press, pp. 32-44.

Fjelland, R. (2020) 'Why general artificial intelligence will not be realized', *Humanities and Social Sciences Communications* 7(10). Available at: https://doi.org/10.1057/s41599-020-0494-4

Gaut, B. (2003) 'Creativity and imagination', In B. Gaut and P. Livingston (eds.) *The Creation of Art*, Cambridge: Cambridge University Press, pp. 148-173.

Gaut, B. (2009). 'Creativity and skill', in Krausz, M., Dutton, D., and Bardsley, K. (eds.) *The Idea of Creativity*, Leiden: Brill, pp. 83-103.

Gaut, B. (2010) 'The philosophy of creativity', *Philosophy Compass*, 5, pp. 1034–1046. Available at: http://doi.wiley.com/10.1111/j.1747-9991.2010.00351.xGaut, B. (2012) 'Creativity and rationality', *The Journal of Aesthetics and Art Criticism,* 70, pp. 259-270. Available at: https://doi.org/10.1111/j.1540-6245.2012.01518.x

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Benigo, Y. (2014) 'Generative adversarial networks', *arXiv*. Available at: https://doi.org/10.48550/arXiv.1406.2661

Gordon, E. C. (no date) 'Understanding in Epistemology', *The Internet Encyclopedia of Philosophy*. Available at: https://iep.utm.edu/understa/#H1

Grimm, S. (2024) 'Understanding', In Edward N. Zalta & Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2024 Edition), available at: https://plato.stanford.edu/archives/win2024/entries/understanding/

Harnad, S. (1989) 'Minds, Machines and Searle', *Journal of Experimental and Theoretical Artificial Intelligence*, 1, pp. 5–25.

Helliwell, A. C. (2021) 'Darwinian creativity as a model for computational creativity', *Proceedings of the 7th Computational Creativity Symposium at AISB 2021 (CC2021),* pp. 15-19. Available at: https://aisb.org.uk/wp-content/uploads/2021/04/cc_aisb_proc.pdf (Accessed: 1 December 2022)

Helliwell, A. C. (forthcoming) 'Creativity, Agency, and AI', In Müller, Vincent C.; Leonard Dung, Guido Löhr and Aliya Rumana, (eds.) *Philosophy of Artificial Intelligence: The State of the Art.* Synthese Library, Springer. (In Press)

IBM (2024) 'What is black box AI?' *IBM Think*, October 24. Available at: https://www.ibm.com/think/topics/black-box-ai (accessed 12 August 2025)

Hills, A. (2009) 'Moral Testimony and Moral Epistemology.' *Ethics,* 120. pp, 94-127.

Hutto, Daniel D. and Ian Robertson, 2020, 'Clarifying the Character of Habits: Understanding What and How They Explain', in *Habits: Pragmatist Approaches from Cognitive Science, Neuroscience, and Social Theory*, Fausto Caruana and Italo Testa (eds.), Cambridge: Cambridge University Press, 204–222. doi:10.1017/9781108682312.010

Howell, N. R. (2007) 'Uniqueness in context', *American Journal of Theology & Philosophy,* 28(3), pp. 364–77. Available at: http://www.jstor.org/stable/27944418

Levinson, J. (2003) 'Elster on artistic creativity', in B. Gaut and P. Livingston (eds.) *The Creation of Art: New Essays in Philosophical Aesthetics,* Cambridge: Cambridge University Press, pp. 235-256.

Livingston, P. (2009). 'Poincaré's "Delicate Sieve": On creativity in the arts', in Krausz, M. D. Dutton, D. and Bardsley, K. (eds.) *The Idea of Creativity*, Leiden: Brill, pp. 129-146.

Luty, J. (2017) 'Do Animals Make Art or the Evolutionary Continuity of Species: A Case for Uniqueness', *AVANT*, 8(1), pp. 107-116.

Michalski, R.S., Carbonell, J.G. and Mitchell, T.M. eds., (1983) *Machine learning: An artificial intelligence approach*. Palo Alto, CA: Tioga Publishing Co.

Moruzzi, C. (2021) 'Measuring creativity: an account of natural and artificial creativity.' *European Journal of Philosophy of Science,* 11(1). Available at: https://doi.org/10.1007/s13194-020-00313-w

Moruzzi, C. (2022), 'Creative Agents: Rethinking Agency and Creativity in Human and Artificial Systems.' *Journal of Aesthetics and Phenomenology* 9(2), pp. 245-268. doi: 10.1080/20539320.2022.2150470.

Moruzzi, Caterina (2025). 'Artificial Intelligence and Creativity', *Philosophy Compass* 20(3), e70030.

Novitz, D. (1999) 'Creativity and constraint', *Australasian Journal of Philosophy*, 77 (1), pp. 67-82. Available at: https://doi.org/10.1080/00048409912348811

Paul, E. S. and Stokes, D (2018) 'Attributing Creativity', In M. Kieran and B. Gaut (Eds.), *Creativity and Philosophy*. Oxford: Routledge, pp. 193-210.

Paul, E. S. and Stokes, D. (2024) 'Creativity', in Edward N. Zalta & Uri Nodelman (eds.)*The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition) available at: https://plato.stanford.edu/archives/spr2024/entries/creativity/.

Pavese, C. (2022) 'Knowledge How', in Edward N. Zalta & Uri Nodelman (eds.) *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), available at: https://plato.stanford.edu/archives/fall2022/entries/knowledge-how

Robertson, I., Hutto, D.D. Against intellectualism about skill. *Synthese* **201**, 143 (2023). https://doi.org/10.1007/s11229-023-04096-4

Russell, S. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. 3rd edn. Upper Saddle River, NJ: Prentice Hall.

Searle, J. R. (1980) 'Minds, brains, and programs', *Behavioral and Brain Sciences,* 3(3), pp. 417–424. Available at: doi:10.1017/S0140525X00005756

Searle J. R. (1999) 'The Chinese Room', in R.A. Wilson and F. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences,* Cambridge, MA: MIT Press.

Sejnowski, T.J. (2018) *The deep learning revolution*. Cambridge, MA: MIT press.

Simonton, D. K. (1999) 'Creativity as blind variation and selective retention: Is the creative process Darwinian?', *Psychological Inquiry*, pp. 309-328.

Sprevak, M. (2007) 'Chinese rooms and program portability', *British Journal for the Philosophy of Science,* 58(4), pp. 755–776.

Stanley, J. and Williamson, T. (2017) 'Skill.' *Noûs*, 51(4) pp. 713–26. Available at: http://www.jstor.org/stable/26631463.

Stravinsky, I. (1997) 'Poetics of music', in Barron, F., Montuori, A., and Barron A. (eds.) *Creators on Creating: Awakening and Cultivating the Imaginative Mind.* New York: Putnam, pp. 189-194.

Stokes, P. D. (2005) *Creativity from Constraints: The Psychology of Breakthrough,* New York: Springer.

Symes, C. (1999) 'Writing by numbers: OuLiPo and the creativity of constraints', *Mosaic: a Journal for the Interdisciplinary Study of Literature,* 32(3), pp. 87-107.

Acknowledgements

I would like to thank Michael Newall, Hans Maes, and the anonymous reviewers for their thorough engagement with this article. Thanks too to the audience at the British Society of Aesthetics annual conference who offered helpful feedback on this paper.