

University of Tartu  
Institute of Philosophy and Semiotics

**CAN COMPUTERS BE CREATIVE?**

Master's Thesis in Philosophy

Patrick Keenan

Supervisor: Eve Kitsik

Tartu 2018

## Table of Contents

<b>1. INTRODUCTION.....</b>	<b>3</b>
<b>1.1 Main aims and claims .....</b>	<b>3</b>
<b>1.2 Some terminology .....</b>	<b>4</b>
<b>1.3 Motivation.....</b>	<b>4</b>
<b>1.4 Outlining the thesis .....</b>	<b>6</b>
<b>2. DEFINING CREATIVITY .....</b>	<b>9</b>
<b>2.1 Definitions of creativity.....</b>	<b>9</b>
<b>2.2 Creativity of agent, process and product .....</b>	<b>12</b>
<b>2.3 Understanding the novelty condition of creativity .....</b>	<b>14</b>
<b>2.4 Understanding the value condition .....</b>	<b>16</b>
<b>3. CREATIVITY REQUIRES CONSCIOUSNESS.....</b>	<b>19</b>
<b>3.1 The ‘what-it-is-likeness’ of subjective experience .....</b>	<b>20</b>
<i>3.1.1 The experiential account .....</i>	<i>20</i>
<i>3.1.2 The jazz experiment .....</i>	<i>22</i>
<b>3.2 The role of imagination in creativity .....</b>	<b>23</b>
<b>3.3 The passive view of creativity.....</b>	<b>24</b>
<i>3.3.1 Traditional passive views.....</i>	<i>25</i>
<i>3.3.2 Freud’s psychodynamic theory .....</i>	<i>26</i>
<i>3.3.3 Bargh &amp; Chartrand’s view .....</i>	<i>27</i>
<b>3.4 Berys Gaut’s agency conception of creativity by flair.....</b>	<b>28</b>
<i>3.4.1 Creative actions must exhibit purposefulness.....</i>	<i>29</i>
<i>3.4.2 Creative actions exhibit a degree of understanding or skill .....</i>	<i>30</i>
<i>3.4.3 Creative actions must exhibit a degree of judgment .....</i>	<i>30</i>
<i>3.4.4 Creative actions must exhibit an evaluative ability directed to the task at hand</i>	<i>31</i>
<b>4. COMPUTERS CANNOT BE CONSCIOUS .....</b>	<b>33</b>
<b>4.1 Searle and the Chinese room.....</b>	<b>33</b>
<b>4.2 Daniel Dennett and the Turing Test .....</b>	<b>35</b>
<b>4.3 Further objections to the second premise .....</b>	<b>36</b>
<b>5. CONCLUSION .....</b>	<b>39</b>
<b>REFERENCES .....</b>	<b>41</b>

# 1. INTRODUCTION

Machines that replace physical labor have allowed us to focus more on what makes us human: our minds. Intelligent machines will continue that process, taking over the more menial aspects of cognition and elevating our mental lives toward creativity, curiosity, beauty, and joy. These are what truly make us human, not any particular activity or skill like swinging a hammer - or even playing chess. (Kasparov 2017, 10)

## 1.1 Main aims and claims

The central aim of this thesis is to investigate and argue against the claim that consciousness is required for creativity. I address this claim, because it is important for answering the question of whether computers can be creative.

There is a core sense of “creativity”, such that creative actions are carried out by creative agents, whose creative achievements we are justified in celebrating and appreciating. This is different than appreciating the beauty of a sunset, or marveling at all the variety of forms of life on earth brought about by evolution. These natural processes might be “creative”, but in a different sense than I will use here. When asking if computers can be creative, we should employ the core sense of “creativity”, such that creativity is the work of agents and not solely the result of natural processes, luck or mechanical operation. Given this core sense of “creativity”, then, creativity is agency-driven, highly valuable and praised in society; and further, it appears in response to a problem or challenge, across domains.

However, what does this core sense of “creativity” imply for the role of consciousness in creativity: does such creativity require consciousness? This is important to ask, because if creativity requires conscious agents, then a computer must be conscious to be creative; and it is unclear whether computers can be conscious. My thesis is a contribution to the discussion of the question of whether computers can be creative. In order to address this question, I will draw on the philosophical literature on creativity, artificial intelligence and consciousness.

I will structure this thesis around one argument:

- (1) Creativity requires consciousness;
- (2) computers cannot be conscious; therefore,
- (3) computers cannot be creative.

Since this is a valid argument, it is impossible for the conclusion ‘computers cannot be creative’ to be false if the premises are true. To disprove the argument, I need to refute only one of the two premises. I will argue that the first premise is false. Disproving this argument will not prove that computer creativity is possible; it will only show that this particular argument fails to show the impossibility of computer creativity.

## **1.2 Some terminology**

To help the reader better understand the question ‘Can computers be creative?’, I will first define a few terms. I explore “creativity” in greater detail in Chapter 2, so for now I will briefly introduce the standard view of creativity in philosophy, according to which “creativity” is the ability to produce artifacts or ideas that are novel and valuable. “Artifacts” are products or ideas produced by agents. The use of “artifact” or “idea” disqualifies as creative processes those which produce novel or valuable products, but are not the work of an agent. In accordance with this standard view, we tend not to think of things like sunsets as “creative” in this core sense, because we know that the novel and valuable result of scattered light rays hitting the atmosphere is not the product of an agent.

An “agent” is generally thought of as a “being with the capacity to act” (Schlosser 2015). An “act” or “action” is something we genuinely do, as opposed to something that happens to us. “Acting”, in this sense, implies certain causal properties – that “I” initiate certain behaviors. However, actions like ‘coughing’ must be distinguished from actions like ‘writing a novel’. I initiate both, but only the latter involves “intention” and “intentional actions” (ibid.). Chapter 3 will elaborate on this distinction.

Finally, I will define how I am using the term “computer”. A modern understanding of “computer” is that it is

an electronic device (or system of devices) which is used to store, manipulate, and communicate information, perform complex calculations, or control or regulate other devices or machines, and is capable of receiving information (data) and of processing it in accordance with variable procedural instructions (programs or software). (“Computer”, *Oxford English Dictionary*)

## **1.3 Motivation**

My aim in this section is to show why the question about computer creativity is worth thinking about philosophically. I will also briefly address some objections to the proposed motivation.

Increasingly, computers are doing the work that humans used to do. This move towards automation is justified by the evidence that computers generally perform these tasks better and are therefore safer, more efficient, more accurate, etc. Furthermore, it is typically thought that computer automation will allow people the freedom to pursue more creative activity, to add more value to their life. Being creative, in turn, is considered crucial to living a good life in that it increases one's sense of self-worth. But what if computers themselves would become creative? Creative computers would be more likely to resolve many of the problems they encounter. Thus it seems natural to assume that computer scientists and programmers would continue their attempts to model and formalize creativity. However, if creativity is essential for human self-worth and sense of purpose, then creative computers may negatively impact humanity.

To better understand this claim, and consider an objection, we can look at the cases of Chess and Go, where two computers, Deep Blue and AlphaGo, beat the reigning world champions in their respective games. Chess and Go were, for a long time, considered classic examples of human intellectual achievement. Before Chess champion Garry Kasparov lost to the Chess supercomputer Deep Blue in 1996, he claimed he was playing to help “defend our dignity” (Wright 1996, *Time*). It was much the same scene twenty years later when world Go champion Lee Sedol lost to Google DeepMind's AlphaGo (a Go-playing supercomputer):

Lee Sedol openly apologized to the Korean public and the wider Go community (and, perhaps, humans in general) for losing the match, tapping an undeniable melancholy among those gathered to watch the match inside Seoul's Four Seasons hotel. (Metz 2017, *Wired*)

Computers have surpassed human abilities in these games, but this does not threaten our self-worth. I think this can be explained by the fact that Chess and Go are games, not problems to be solved. As games, they are fun to play, so people will continue to play them despite computer capabilities. At first glance, the same line of thinking could apply for creativity, as well: It is fun to be creative, therefore we will continue to be creative, even if computers are creative.

The difference is that creativity is a response to a situation, i.e. a ‘creative solution’. People are creative in game-playing, but the game is not dependent (for its existence) on not being solved. In this regard, I have two assumptions: (1) the core sense of agency driven creativity is the same concept across different domains; and (2) that creativity is always a case of problem-solving. I will explain (1) in Chapter 2 when I make the distinction between the subjective creativity of the agent and the objective creativity of artifacts or ideas. Regarding (2), aestheticians might object: “the artist, unlike the scientist, is not standardly confronted with problems to solve” (Gaut 2010, 1038). This is the appeal of *art pour l’art*, that “under the sun there neither exists nor can exist any work more thoroughly dignified, more supremely noble, than this poem which is a poem and nothing more, this poem written solely for the poem’s sake” (Poe 1850, 5). The expression “for the poem’s sake” works because artists value their art, Poe even appeals to the nobility of the act. But one could make the same argument for murder, ‘murder for murder’s sake’. I would say that artists value art, therefore art is worth pursuing, and so creating art has a purpose. Nietzsche says “does his basic instinct aim at art, or rather at the sense of art, at life? at a desirability of life? Art is the great stimulus to life: how could one understand it as purposeless, as aimless, as *l’art pour l’art*?” (Nietzsche 1889. §24, 80). If creativity is always a response to a problem, and computers will be tasked with solving problems across domains, it is safe to assume that human creativity could be superseded by computer creativity.

In this section, I showed how creative computers could challenge human self-worth. Humans add value to their lives by being creative. Creativity is a valuable problem-solving resource and so there are strong incentives to make computers creative. Computers are already taking over many human tasks and outperform humans in more and more diverse domains. Perhaps creative computers could even take over all problem-solving. If this happens, it will be hard to know how human beings can continue to add value to their lives. In this way, human self-worth is challenged by creative computers.

#### **1.4 Outlining the thesis**

Thus far, I have presented the argument I am looking at and I have stated my main aim, which is to argue against the first premise, i.e. that creativity requires consciousness. I have also shown why the question about computer creativity is worth

considering (because creative computers may challenge human self-worth). I will spend the bulk of the rest of my thesis on creativity, first defining “creativity” (Chapter 2) and then addressing the first premise of the master argument: ‘Creativity requires consciousness’ (Chapter 3). Chapter 4 will deal with the second premise: ‘Computers cannot be conscious’. In Chapter 5, I will provide some concluding remarks.

More precisely, in the second chapter, I will discuss the standard view of creativity and address some issues with the “novelty” and “value” conditions of the standard view. I will consider “creativity” to be applicable across all domains of human activity, so I will briefly explain why science should be treated as a creative enterprise (against Kant). I will distinguish between subjective and objective creativity and defend my focus on subjective creativity: the psychological aspect of creativity (which involves the experiences of the creator). Finally, I will try to rehabilitate the standard view against the arguments of Hills and Bird, according to which creativity need not include the value condition.

In the third chapter, I will assess the first premise of the consciousness argument, that creativity requires consciousness. I will look at traditional views of creativity that Berys Gaut classifies under “passive creativity”: views according to which creative acts do not result from the agent’s conscious efforts. The consideration of these views motivates turning to contemporary literature on the role of the unconscious. These contemporary views might help explain the origins of creative ideas without appealing to gods or madness. “Passive creativity” is contrasted with “active creativity”, which involves conscious deliberation in the use of imagination, which, I will show, is not necessary for creativity. I will also address the four conditions that Berys Gaut lays out in his agency-view of creativity, which involves the standard view (the novelty and value conditions), plus “flair”. “Flair” is shorthand for purposefulness, understanding, judgment and evaluative ability. I will ask whether these conditions are necessary for creativity and whether they entail consciousness.

In the fourth chapter, I will assess the second premise of the consciousness argument – that computers cannot be conscious – by looking at John Searle’s argument that a functional simulation of a conscious state like ‘understanding’ does not qualify as *real* understanding. I will assess Paul and Patricia Churchland’s response to Searle, along with their dismissal of the Turing Test as a measure of conscious intelligence. I will then look at Daniel Dennett’s defense of the Turing Test.

Finally, I will address several arguments for and against conscious computers from Daniel Dennett and Hilary Putnam.

I will conclude my thesis with a brief summary of the ideas I have discussed and delineate concluding remarks outlining the specific role (if any) of consciousness in creativity and whether or not it could be instantiated in computers.

## **2. DEFINING CREATIVITY**

In this chapter, I will build towards a working definition of creativity. To begin with, I will look at how other philosophers have defined creativity. Then, I will characterize creativity as consisting of three necessary parts: agent, process and product. It is safe to say that these distinctions underlie all examples of creativity. Next, I will analyze a recurring theme in descriptions of the creative process: the distinction between “generation” and “selection” as phases of the creative process. Similarly, I think that this distinction applies to all examples of creativity. Finally, I will investigate both ‘novelty’ and ‘value’, as the standard necessary conditions in definitions of creativity, and point to issues related to each one. I conclude the chapter by introducing a modification of the standard view. According to that modification, creativity aims at value, though it does not always succeed. The overall account of creativity I give emphasizes the role of ‘taste’ as the active ingredient in all relevant cases of creativity. Taste seems to involve consciousness; namely the subjective experience of ‘qualia’ or phenomenal consciousness (that there is something it is like to enjoy x). Therefore, this conclusion will pave the way for exploring the role of consciousness in creativity in Chapter 3.

### **2.1 Definitions of creativity**

The standard view of creativity is that it is the ability to produce novel and valuable ideas or artifacts. The first chapter emphasized the core, intuitive sense of creativity, which involves ‘artifacts’ as created by beings with a capacity to act, as opposed to ‘objects’ created by natural processes or luck. Most philosophers define creativity with variations on those central conditions of ‘novelty’ or ‘value’, each with their own nuance and particular emphasis on one or both, along with a variety of supporting features. I will point out some of those supplemental features when looking at different philosophers’ conceptions of creativity. For now, it is helpful to think of creativity as I have defined it above; but note that the conditions of novelty and value are not without philosophical contention, which I will highlight later in this chapter.

Margaret Boden, who exemplifies the standard view, defines creativity as “the ability to come up with ideas or artifacts that are new, surprising and valuable” (Boden

2010, 29). For Boden, ideas and artifacts include poems, musical compositions, cooking recipes, jokes, paintings, origami, penny whistles, steam engines, etc. (ibid.). Given the many tokens of creativity on display here, it is apparent that creativity features in nearly every aspect of life. Boden even says creativity is “an aspect of human intelligence in general” (ibid.). Because creativity, so understood, is a stable feature of human intelligence, she states that “every one of us is creative, to a degree” (ibid., 30). Maria Kronfeldner’s view is another example of the standard view. She cites Boden’s definition of creativity and also Sternberg and Lubart’s definition, according to which “creativity is the ability to produce work that is both novel (i.e., original, unexpected) and appropriate (i.e., useful, adaptive concerning task constraints)” (Kronfeldner 2009, 578). Kronfeldner’s own definition of creativity is that “creativity, in its basic sense, is the human capacity to create new and valuable responses to challenges to which humans are exposed, or to which they expose themselves” (Kronfeldner 2014, vii).

Both Boden and Kronfeldner see creativity as existing in every domain. Immanuel Kant, however, distinguishes the sciences from the arts, specifically fine art, and ties his notion of genius (creativity) only to works of fine art, but not to the sciences (Kant 1790/1987, §46, §47). The relevant sense of “creativity” here applies to all domains; so I must argue against Kant regarding creativity in the sciences. Let me first present his reasoning, and in doing so, I will give Kant’s definition of creativity as well. According to Kant, creativity is exhibited by genius, and genius is “the innate mental predisposition through which nature gives the rule to art” (ibid., 174). What Kant is saying, more precisely, is that creativity is a ‘natural talent’ which exists in the agent’s ability to communicate to the general audience via ‘the understanding’ the universally shared expressions of beauty and sublimity present in nature, through the creation of an object of fine art. Fine art is differentiated from craft or the lighter agreeable art (art simply for pleasure), in that fine art not only appeals to pleasures but is cognitively stimulating, by exciting in the audience the same spirit which the artist had when creating.

The mental predisposition that is sufficient for creating art cannot just be expressing oneself in a particular artistic medium; for not all such products exhibit good taste. Likewise, applying good taste is not sufficient for creating art; if that were the case, then any art critic would be an artist. Mediocre products lack spirit, which

Kant calls the “animating principle of the mind” (ibid §49, 181). He considers the excitations of the audience to mimic the experience of creating. Accordingly, if a work of fine art is the product of genius, then the audience will find it both pleasurable and cognitively stimulating. Not only does Kant presuppose a mental process in creativity, but he also distinguishes between two aspects of the creative process: generation and selection. “Genius provides the ‘matter’ for fine art (generation), taste provides the ‘form’” (selection) (Burnham IEP Immanuel Kant: Aesthetics).

Kant’s definition of creativity departs from the assumption that the genius creates in such a way that no determinate rule can be given for it. Therefore “its foremost property is originality”; but “since nonsense too can be original”, products of genius must also be exemplary (Kant 1987 §46, 175). So here we have two very similar conditions to those given above by Boden and Kronfeldner: the products must be original (the novelty condition) and exemplary (the value condition). But Kant adds something to the standard view when he claims that the genius cannot know or explain how their ideas came about. This allows him to argue that scientists are not creative. “Genius must be considered the very opposite of a spirit of imitation” (Kant 1987 §47, 176) and if learning is imitation, then great scientific discoveries could be learned as well; this is the “natural path of an investigation and meditation by rules” (ibid.). Kant says one can learn everything Newton put forward in his principles of natural philosophy because Newton can show how every one of his steps came together from start to finish. But no “Homer or Wieland can show how his ideas, rich in fancy and yet also in thought, arise and meet in his mind; the reason is that he himself does not know and hence also cannot teach it to anyone else” (ibid., 177).

Because I consider “creativity” applicable to the sciences as well as the arts, I need to challenge Kant’s view. Ted Cohen puts forward a strong case against Kant’s reasoning. Kant claims that Newton’s work is entirely explicable in terms of concepts. However, Cohen states that despite the fact that “Newton’s Principia is explicable conceptually, nothing whatsoever follows about how, in fact, Newton produced the work” (Cohen 2003, 140). Cohen emphasizes that Newton’s presentation of his discovery “may or may not mirror his own initial reasoning” (ibid.). In other words, the products of scientific labor tell us nothing about how the ideas came to be.

Cohen gives an example from his own experience in which he was tasked with trying to understand one of Gödel’s Incompleteness Theorems. In coming to a full

understanding, Cohen explored various proofs from professors and other writers, but had not actually engaged Gödel's original texts. When Cohen had a strong understanding of the First Incompleteness Theorem, he inferred that Gödel himself probably went through the same steps that Cohen had. Later, when he approached the source texts, he was "surprised at some of the features of Gödel's own exposition" and found that he had even less of an idea how Gödel himself came to his conclusions. Cohen's conclusion, which I agree with, is

...that it is an error to suppose that the discovery of these theses proceeded as proofs do. And it is still a mistake to suppose that the author must have followed some particular series of steps in arriving at his discovery and that these steps must be capable of being articulated conceptually (ibid.).

As the next chapter will show, there are many scientific discoveries which appear completely unlike the logical entailment of one idea to the next, and many introspective reports even border on the mystical. The "magic" of the initial idea is the real mystery and that mystery seems to be there both in the arts and in the sciences.

In this section, I have presented the standard view of creativity and looked at three different philosophers' accounts of creativity: Margaret Boden, Maria Kronfeldner, and Immanuel Kant. I have shown how the concepts of "novelty" and "value" are staples of all these definitions. I also discussed Kant's argument against creativity in the sciences and found that, following Cohen, we should consider creativity as existing in the sciences as well. In the next section, I will look at creativity in terms of the agent, the process and the product.

## **2.2 Creativity of agent, process and product**

"The term 'creative' is used to describe three kinds of things: a person, a process or activity, or a product, whether it is an idea in someone's mind or an observable performance or artifact" (Kaufman, S., & Paul, B. 2014, 6). There is disagreement about which of these three aspects (if any) are fundamental to creativity. If creativity is the same concept across different domains, there must be something essential to creativity which can be represented in each domain. This section will look at each of these parts and defend focusing on the creator as the fundamental aspect of the three.

“Objective creativity” focuses on the artifacts of creative achievement. Looking at the artifacts cannot tell us if creativity has taken place. For example, the formula for mass-energy equivalence - “ $E=mc^2$ ” - does not tell us anything about how Einstein came up with his formula. It is possible he discovered it entirely by accident. This would not change the formula’s value, but it would be more difficult to claim that Einstein was creative in his discovery.

Next, we will look at the creative process. Normally, being creative is often considered the opposite of being mechanical, so calling any mechanical process “creative” seems to violate this core intuition about creativity. However, creativity may be employed by mechanical means. An assembly line, for example, produces a car and many different people play a role in this production. Each person has a very specific repetitive task and the final product is a new car, surely valuable, and an artifact made by agents. To find the core sense of creativity, we have to look at the original designer who drafted the plans from scratch. The ‘artifact’ may be realized mechanically, but creativity is still present. In many cases, it is actually impossible to realize a creative idea without employing some kind of mechanical process. Simply evaluating the process could cause us to reject cases of creativity which can only be realized by mechanical means. To understand the core, intuitive sense of creativity, we have to see how the idea was itself generated.

Subjective creativity emphasizes the agent and how their ideas come into being. This distinction is important to the overarching discussion about computers and their products. Presently, we have creative artifacts made by computers, which even employ a process that replicates the psychological mechanisms of human creativity. Margaret Boden calls these results “apparent” creativity, distinguishing them from “genuine” creativity, the kind exhibited by human beings (Boden 2014, 225-6). Boden shows how creative psychological processes could be instantiated in computers, by distinguishing between three types of creativity that emerge from different psychological mechanisms: combinatorial (the juxtaposition of diverse ideas), exploratory (novel structures informed and limited by their defining constraints) and transformational, or “radical creativity”, which explores the conceptual space and rejects aspects of the framework in favor of a new paradigm (Boden 2014, 228). Computationally instantiating these processes lead to novel and valuable results without necessarily being “genuine”, whether that means involving conscious mental states or not. On one

hand, it could be claimed that whether they have mental states or not, computers can be creative, if they instantiate the relevant processes. But, on the other hand, we could dismiss computer creativity by asserting that no amount of computer creativity could ever produce something like Kant's fine art and that being "genuine" in the core sense of creativity means a computer would have to ascend to the heights of Chopin or Shakespeare – something that we do not even require of human beings whom we call "creative". We need to find a less demanding condition that separates apparent creativity from genuine creativity, or mere art from fine art. Kant says it is taste that "consists in disciplining (or training) genius" (Kant 1790/1987 §50, 188). Without taste, or judgment, all of the imagination's "riches (in ideas) produce nothing but nonsense" (ibid.).

Art can be distinguished from fine art with the latter exhibiting the productive capacity of genius shaped by judgments of taste. Psychological mechanisms for creativity can be instantiated into computers which, by virtue of their design can exhibit the necessary skills, producing the content of genius; however, judgments of taste might be out of bounds for computers, because they would require subjective experience. If so, then computers would lack the capacity for "genuine" creativity. But this claim only holds if: (a) taste is necessary for creativity and necessarily conscious, and (b) computers cannot be conscious.

In the next two sections, I clarify the conditions of novelty and value.

### **2.3 Understanding the novelty condition of creativity**

In the last section, the importance of looking to the subjective experience of creativity was emphasized. In connection with this, Margaret Boden puts forward a relevant clarification regarding novelty. She distinguishes between "psychological" and "historical" creativity (Boden 2010, 30-31). P-creativity (psychological) refers to a creative idea that is new to the person creating it. H-creativity (historical) refers to an idea that no one (in history) has ever had before.

This is important to distinguish because, in the case of someone being P-creative, they are still being creative even if it turns out that their creative idea already exists in the world. Charles Darwin and Alfred Russel Wallace are a classic example of this, since both put forward a theory of evolution through natural selection almost

simultaneously without having had access to each other's work. It would seem that novelty implies a kind of newness which (taken too strictly) would invalidate the creativity of someone – either Darwin or Wallace – based on who published first. But this is contrary to our intuitions about creativity. And this is why P-creativity is a necessary condition for creativity. Anything that is H-creative will also be P-creative. H-creativity is the concern of objective creativity, and so I will not be considering it in depth in this thesis.

Novelty can be trivial. In the broadest sense of anthropological, or cultural novelty, culturally new ideas, patterns and behavior transmitted through generations count as “novel”. “In this anthropological sense, creativity is a *conditio humana* and a ubiquitous activity” (Kronfeldner 2009, 579). In this most basic sense, creativity is a natural part of being (this even echoes Boden's claim that creativity is a basic feature of human intelligence). These everyday acts represent a subjective, yet banal creativity: a creativity that is ordinary. Robert Weisberg gives philosophical support to the idea that novelty is the norm of all behavior, “because we are continually adapting our behavior to fit new circumstances” (Weisberg 1993, 11). It is normal to produce objects or ideas of minimal value that are novel, such as a salad for dinner, or precisely the way in which I walk to school. The ubiquity of (trivial) novelty – that it might even be impossible not to be novel – puts all the onus for distinguishing creative actions from other actions on value.

One way to resolve this issue might be by considering one final variety of novelty that Kronfeldner describes, namely metaphysical novelty. If the universe is a naturalist, determinist kind, the only kind of novelty that could then be genuine would have to come to the mind of the creator independent of any causal determination. On the deterministic view, all novelty is fundamentally reducible to antecedent pre-conditions. Some philosophers argue that genuine creativity must feature *creatio ex nihilo* (creativity out of nothing), otherwise any newness would simply be the consequent of antecedent conditions (Hausman & Rothenberg 1976, 3-26). In a naturalist universe, metaphysical novelty is impossible because the natural world is all there is; in a determinist universe, metaphysical novelty would be the only genuine novelty because any novelty that did not come from a metaphysical realm would have to have been predetermined. This is also known as the paradox of creativity: “genuine creativity is either unexplainable or not present” (Kronfeldner 2009, 582). I think we

can safely assume a naturalist universe, and so explanations for novelty should conform to the naturalist paradigm. If anthropological novelty is too broad a category and trivializes novelty, then metaphysical novelty is too strict and makes genuine novelty near impossible to obtain.

This discussion of novelty seems to place a much heavier burden on the role of value in creativity. In this next section I will consider Alison Hills' and Alexander Bird's argument that it is possible to be creative and not create anything of value. Their definition of creativity "would reject value but retain originality" (Hills & Bird 2018, 95). This will be difficult to resolve with the emphasis I have placed on value. To redeem novelty from triviality, we continue to value psychological novelty (newness to the creator) of a certain kind and may need to adjust the standard view to accommodate the argument for creativity without value.

## **2.4 Understanding the value condition**

Allison Hills and Alexander Bird argue that it is possible to be creative without producing objects of value. In their view, creativity includes novelty (originality), imagination, fertility and motivation, but not value. In this section, I will address their argument against value. In the previous section, I discussed the ubiquity of novelty and its trivialness as simply a natural part of behavior. On this account, the role of value was emphasized. It would be fatal to the standard view if novelty were treated as inseparable from ordinary behavior and value had no role. Hills and Bird's argument against the value condition may prove difficult to reconcile with the importance of the value condition for addressing the problem with the triviality of novelty. But it is possible that we can adjust the standard view to accommodate their argument in a way that keeps both novelty and value in the definition of creativity.

Hills and Bird's main definition of creativity is that it is the "disposition or set of linked dispositions of an individual: to have many ideas (fertility); which are novel (originality) and generated through use of the imagination (imagination); and to carry through these ideas to completion (motivation)" (Hills & Bird, 2018, 95). Their definition lacks a value requirement. If creativity requires value, it should be impossible to be creative and produce objects with no value.

There are many ways in which value can be obtained in a creative act. An idea or artifact can: (1) have objective value; (2) be good of its kind; or (3) be good for the creator (Hills & Bird 2018, 96-97). The core sense of “creativity” still counts as “creative” cases which can be shown to meet none of the above conditions of value. In other words, even when value cannot be found, we intuitively think creativity has happened. Without introducing supplemental conditions, is there any way to accommodate such cases and keep the value condition?

We can look at Hills and Bird’s case of the failed torturer. Torture instruments, in general, have no objective value (ex. 1); and this particular torture instrument is a failed variation of ‘the rack’, e.g. “it causes death too quickly, without enough suffering” (ibid., 98) and so is not even good ‘of its kind’ (ex. 2); finally, producing torture instruments makes the torturer miserable (ex. 3). Do we still consider the torturer creative? If yes, how can this be if there is no value to be found anywhere in the above case?

Given Hills and Bird’s reasoning, I believe that the standard view could be tweaked in a way that still includes a value condition. According to the modified standard view, being creative always aims at producing something that is valuable. Thus, creativity involves an agent aiming to increase their idea of value. An objection here would be that someone could be forced to create a torture instrument, even though they are against torture. In that scenario, their actual goal is to make the most dysfunctional torture instrument. That is the context in which they are aiming at value. Instead, if their life depends on making the ‘best’ torture instrument, they will aim at that value if they value their life more. Assuming the torturer is not aiming to fail at making torture devices, he is still being creative. ‘Aiming’ points to an attempt at value. I think it is fair to state that no one who we consider creative aims to produce ideas or artifacts of *un*value. Rather they aim at value, and succeed or fail. But success or failure of an act should not invalidate the creativity.

Another objection to my claim might be that ‘aiming at value’ is not the same as actually having value. I think this is the principal distinction between subjective and objective creativity. Subjective creativity aims at value; objective creativity succeeds. I think this is partially addressed by Hills and Bird’s argument that the creative disposition is manifestly the same, regardless of the outcome. So, one is creative even if what one produces has no value. And one is not being creative if one aims at *un*value.

The torturer's aim is certainly to create something of value in the context of torture instruments. He succeeds or fails, but his creativity aims at value, regardless.

Hills and Bird could be read as echoing this claim when they acknowledge that success or failure of creativity does not signal two different dispositions: "one to imaginatively produce things of value; one to imaginatively produce worthless things" (ibid 97). The same disposition aims at value and subjective creativity should not be measured by success or failure in an objective sense. This is also why Hills and Bird put forth imagination as their most important supplemental characteristic of creativity. Imagination does not guarantee success, but the more vivid one's imagination, the more likely one is to produce highly creative artifacts.

In this chapter, I have discussed the standard view of creativity in philosophy which is that creativity is the ability to produce novel and valuable artifacts. Highly valuable products and artifacts exist which come about through mechanical means or luck, both of which are considered incompatible with the core sense of creativity. A deeper, more intuitive sense of creativity looks to the psychological processes of the agent. With novelty, there is a trivial novelty which human behavior in everyday circumstances automatically produces. This places more of an emphasis on the value condition. However, Hills and Bird have argued that it is possible to be creative without producing anything of value. To salvage the value condition, a modified version of the standard view would claim that creative agents aim at value, not *unvalue*.

### 3. CREATIVITY REQUIRES CONSCIOUSNESS

Artists have a vested interest in our believing in the flash of revelation, the so-called inspiration... shining down from heavens as a ray of grace. In reality, the imagination of the good artist or thinker produces continuously good, mediocre or bad things, but his judgment, trained and sharpened to a fine point, rejects, selects, connects... All great artists and thinkers are great workers, indefatigable not only in inventing, but also in rejecting, sifting, transforming, ordering. (Nietzsche 1878, quoted in Popova 2012)

In the previous chapter, I gave a working definition of creativity. The standard view of creativity is that it is the ability to produce novel and valuable artifacts or ideas. Issues with novelty and value were discussed. Given my motivation for asking if computers can be creative (namely, that computers might become “creative” in the sense in which human creativity is valuable and meaning-giving), it makes sense to focus on P-creativity, or psychological novelty. In P-creativity, an idea is new to the originator.

This chapter focuses on the first premise of the consciousness argument: that creativity requires consciousness. Once we adopt the view of creativity as a subjective phenomenon (“subjective creativity”), outlined above, it seems that the subjective experience of consciousness may play an inescapable role in creativity in connection with taste (the selection phase). Taste was thought to explain the distinction between art and fine art, according to Kant, and here it can be thought to distinguish ‘apparent’ from ‘genuine’ creativity of computers (Boden 2014, 225-6). However, it was also pointed out that creativity cannot simply consist in having taste, for then any art critic would be an artist. Even if taste were necessarily conscious, it is not clear that taste is necessary for creativity. Section 3.1 will look at the experience of taste through Bence Nanay’s experiential account, followed by a brief view of “phenomenal consciousness” according to Thomas Nagel. Doing this will show what an experience of taste must consist in, namely there is something it is like to experience something as not thought to be possible before.

It is also not clear what role consciousness plays in the generation (production) phase of creativity. To address that issue, I will look at the role of imagination and how two views of creativity – passive and active – have implications for the role of consciousness in creativity (3.2, 3.3). Finally, I will discuss four conditions that a creative action must supposedly exhibit. Berys Gaut calls these conditions collectively “flair”. They include purposefulness, understanding, judgment and evaluative ability

(3.4). I will ask if these conditions are necessary to creativity and whether or not they require consciousness. Considering Gaut's view will help with assessing the claim that creativity requires consciousness.

### **3.1 The 'what-it-is-likeness' of subjective experience**

The distinction between subjective and objective creativity relates to the question of what the explanatorily basic notion of creativity is. In the last chapter, I defended focusing on creativity as essentially a mental process; therefore, to better understand it, we must look to the mind of the creator. This section will assess the subjective experience of creativity to determine its necessary components.

#### *3.1.1 The experiential account*

Bence Nanay claims that a mental process is creative "if and only if it produces an idea that is veridically experienced as something we have not thought to be possible before and as something we have not learned from someone else" (Nanay 2014, 24). This experiential account of Nanay's helps to explain what creativity requires, beyond trivial novelty, by distinguishing subjectively creative actions from the trivially novel ones. According to this account, I am being P-creative when I experience my idea as something I have not thought possible before. The problem that novelty could be trivial (e.g. a new salad) disappears under Nanay's account, because I will not experience the salad as something I have not thought to be possible before. Can computers be creative on this account? Yes, if they can have (the relevant) experiences.

Presently, it seems that computers lack experiences. One of the issues with this claim is that there might be no way to know objectively what a specific experience is like (if you are not the one having it), or whether computers are having experiences at all. One might say that experience has a subjective character and there is no objectivity of experience. From the outside, one can understand the physics of something like lightning, but there is no way to explain the experience of lightning in physical terms. And the further we go into the physicalist account of what lightning is, the further away we get from the subjective experience of lightning.

What constitutes a conscious experience? Nagel says that "the fact that an organism has conscious experience at all means, basically, that there is something it is like to be that organism" (Nagel 1974, 436). The subjective character of experience is

not analyzable in terms of “any explanatory system of functional states, or intentional states” (ibid.) since “any shift from a specific point of view to one of greater objectivity takes us farther away from the real nature of the phenomenon: our specific viewpoint” (Nagel 1974, 445).

Conscious experiences have a distinctive feel; there is therefore something it is like to have a conscious experience, but not all experiences might be conscious. Peter Carruthers, for example, recognizes that there are nonconscious experiences: “What does it feel like to be the subject of a nonconscious experience? It feels like nothing” (Carruthers 1989, 506). If we experience an idea as something we have not thought to be possible before, however, there is something it is like to have that experience, and that experience is a conscious experience. This means if Nanay’s experiential account adequately captures subjective creativity, then subjective creativity requires consciousness.

Is Nanay’s condition actually necessary for creativity, however? I do not think so. On his account, it would be impossible to be creative without “experiencing an idea as something not thought to be possible before” (Nanay 2014, 24). New ideas are always technically experienced as something not thought to be possible before, even uncreative or nonsensical ideas.

Additionally, creative ideas may be generated without any conscious experience. There are minimal accounts of agency which grant agency to human and non-human agents without the ascription of mental representations. One example is skilled coping, which is “the ability to engage with others and with one’s circumstances by responding to the demands of the situation in a skillful and often effortless manner, without conscious deliberation, reasoning, or planning” (Schlosser 2015). Examples of skilled coping include improvisation, such as jazz playing or constructing sentences in conversation. Minimal views of agency allow for these kinds of habitual actions to qualify as agency, although they do not involve conscious deliberation. Peter Carruthers argues that there is nothing it is like to experience habitual actions (they are nonconscious experiences) and his examples of habitual actions include driving a car and doing the dishes (Carruthers 1989, 505).

It seems counterintuitive to claim that there is nothing it is like to improvise in jazz, for jazz improvisation should be peak creativity. There is also some empirical evidence that instead shows that skilled coping may actually require consciousness.

This is important because until now, I have only suggested that taste (what one likes) is the necessary conscious experience in creativity. On Nanay's experiential account of creativity, combined with Nagel's idea of phenomenal consciousness as the relevant kind of experience, taste occurs after-the-fact of productive creativity (generation). Generation, however, does not seem to necessitate a role for consciousness. The role of consciousness in generation is challenged by claims that improvisation, as habitual action, does not require conscious deliberation, and therefore there is nothing it is like to improvise jazz. In the next section, I will look at the jazz experiment to see if there is any evidence for this claim.

### *3.1.2 The jazz experiment*

Recent psychological experiments seem to confirm the major role that consciousness plays in the generation of creative ideas. Baumeister, et al., ran a series of experiments to test the role of consciousness in creativity (Baumeister et al. 2016). This section will look at one such experiment which involved musical improvisation of jazz guitar under a cognitive load. The idea was to burden the conscious mind with the task of counting backwards in sixes from the number 913 and simultaneously improvise a guitar solo over accompaniment. This test was based on the assumption that "the conscious mind can only do one thing at a time" (ibid., 188) and so occupying the conscious mind with counting backwards would free up the unconscious mind to be creative. The experiment challenges the common belief that "creativity flourishes when unconscious processes are permitted to flow and blossom undisturbed by conscious interference" (ibid., 186). The musicians themselves were under the impression that their playing would be better because "the conscious ego would be out of the way and the creative unconscious would be able to operate unfettered" (ibid., 191).

Contrary to the expectations of the musicians, the actual creativity of their playing suffered. It was shown that the improvised melodies "tended toward repetition and simple, stilted phrases" (ibid.). What is not clear is whether conscious generation or conscious selection was affected by this experiment. The unconscious seems to be the seat of the spontaneous generation of ideas, but unlike the conscious mind, does not get bored with its solutions to problems. "Boredom may be a conscious feeling, something largely alien to the unconscious" (ibid., 192). The authors concluded that "only with full conscious awareness does the mind reject repetition and continue to exchange different patterns until a fully, pleasingly creative result is obtained" (ibid.).

This experiment aimed to establish whether or not there is a clear role for consciousness in creativity. Although the experiment seems to show that consciousness is necessary, it is not evident in what capacity. If creativity stems from the unconscious, then perhaps the imagination is necessary for expression.

### **3.2 The role of imagination in creativity**

This section looks at the role of imagination in creativity. If one is being actively creative, ideas are entertained in the imagination where they can be freely explored through conscious deliberation. Creative ideas can appear as a product of improvisation, but improvisation may not appear as “something which one has not thought to be possible before” or even with phenomenal characteristics. The inability to apply one’s taste might explain why creativity is hampered in improvisation with inhibited consciousness, but the above experiment concentrated on the generation component of creativity, not the selection component. This section will delve deeper into the role of imagination in active creativity and ask if imagination is necessary for creativity.

What is the core sense of imagination, then? Berys Gaut claims that

imagining that such and such is the case, imagining that *p*, is a matter of entertaining the proposition that *p*. Entertaining a proposition is a matter of having it in mind, where having it in mind is a matter of thinking of it in such a way that one is not committed to the proposition’s truth, or indeed to its falsity (Gaut 2003, 152).

Gaut discusses other uses of ‘imagination’ but contends that the core sense of imagination relates to entertaining a proposition without committing to its truth or falsity. Unlike other mental states such as belief and intention, imagination lacks any intrinsic end at all, thus it “exhibits a kind of freedom” (ibid., 160) making it “suited of its nature - to be the vehicle for active creativity” (ibid.). Active creativity seems to necessitate a conscious role - that of active imagination - for creative acts.

Gaut makes a distinction between imagination as a source of creativity, and as a vehicle for creativity. When Kasparov played Deep Blue, both the human and the computer used their imagination. Deep Blue has, in a sense, a powerful imagination in that it can “survey vastly more possible positions than any human could” (ibid., 158). Gaut argues that Kasparov used his imagination creatively. Kasparov imagines different moves he could make, but it does not follow that his imagination “is the

source of that creativity” (159). Instead, Gaut argues that his creativity is displayed in “how he uses his imagination, but that in turn is explained largely by factors such as his vast experience, considerable knowledge of chess history, practiced technique and sheer native talent” (159).

So what is the source of creativity? “Our own experience provides the basic material for our imagination” (Nagel 1974, 439). While this is true with Kasparov, it is also true with Deep Blue whose “experience” is the rich history of chess books and previous games it was given that it later utilized when conducting its strategy. Though the imagination might be perfectly suited to entertain various propositions, the imagination is not necessary for creativity. For example, Bertrand Russell, when writing *Principia Mathematica*, would work on a problem, then fall asleep and awake the next morning with the solution (Gaut 2003 155). Many creative acts happen without using our imagination and not every use of the imagination results in a creative act.

### **3.3 The passive view of creativity**

There is a set of views on creativity that Gaut collectively calls the “passive view” (Gaut 2003, 150). What these views all share is the idea that the creative act originates from outside the conscious mind of the creator. It could be the case that creativity requires both active and passive mental processes. However, if it turns out that genuine cases of creativity exist without any involvement of consciousness, then the first premise of the main argument of this thesis is defeated, and the argument fails. If any of the processes described in the passive views could result in creativity without conscious intervention by the agent, then creativity would not require consciousness. So far, I have considered whether imagination and taste might be conscious elements that are necessary for creativity. (Taste and imagination are a coupling analogous to generation and selection.) Taste, I have argued, seems inseparable from the core, relevant sense of creativity, but there are problems with this which I will attend to later. Imagination was discussed in the last section and I found that it is not necessary for creativity. The question to keep in mind when looking at the passive view is whether or not novel and valuable artifacts can be produced under the passive view, without any role for consciousness.

### 3.3.1 Traditional passive views

For a historical account of the passive view, we can start with the ancient Greeks who believed that creativity was the work of divine inspiration or divine madness. “Homer begins the Iliad by invoking the goddess to sing his story” (Blackburn 2014, 147). Plato, in *Ion* writes: “...there is no invention in him until he has been inspired and is out of his senses...” (ibid). Even Shakespeare also wrote about divine madness, “comparing the poet’s frenzy to that of the lunatic and the lover” (Gaut 2003, 148). This way of thinking about creativity continues into contemporary writers as well. We can see the nod to the divine essence in an excerpt from Bukowski’s poem “So you want to be a writer?”:

when it is truly time,  
and if you have been chosen,  
it will do it by itself  
and it will keep on doing it  
until you die or it dies in you.  
there is no other way.  
and there never was. (Bukowski 2003).

And these cases are not limited to the arts, the sciences have their own cases of the passive view. Friedrich von Kekulé’s discovery of the closed loop structure of the benzene molecule is portrayed by Kekulé himself here:

I turned my chair to the fire and dozed... Again the atoms were gamboling before my eyes. This time the smaller groups kept modestly in the background. My mental eye, rendered more acute by repeated visions of this kind, could now distinguish larger structures, of manifold conformation; long rows, sometimes more closely fitted together; all twining and twisting in snake-like motion. But look! What was that? One of the snakes had seized hold of its own tail, and the form whirled mockingly before my eyes. As if by a flash of lightning I awoke. (Kronfeldner 2009, 587).

Simon Blackburn talks about the moment of sudden illumination encapsulated in the ‘flash of lightning’, or the ‘stroke of genius’ or Archimedes’ famous outburst of discovery: “Eureka!” (Blackburn 2014, 151).

It may have been Kant who reframed the traditional views of creativity and put the more modern emphasis on the imagination, but, even for Kant, there lingered behind a great mystery that fine art was so enchanting that surely it could not have come from a man. Kant grants the genius a quality that cannot be learned or taught, and cannot even be communicated “but must be conferred directly on each person by the hand of nature” (Kant 1790/1987, §47, 177).

It is not clear what role consciousness plays in the passive view. In none of these cases can we outright claim that consciousness plays no role in creativity, but it is not difficult to imagine that these cases require almost no active conscious deliberation. The passive views signal that “the conscious self of the artist was not the source of the creative work but rather received it as a gift” (Baumeister et al 2014, 186).

A problem with these passive views is this. If the artist or scientist was the conscious recipient of the “gift” of creative work, would they still be praiseworthy in any meaningful sense? A part of the core sense of “creativity” that is relevant here is that creative agents are responsible for their creations. It is no help to call someone “creative” if all that it means is they have received a gift. However, if we discard the traditional inspiration and madness views of creativity in history and the genius view of Kant, how could we explain the origin of creative ideas? There may still be some understanding to be gained by exploring the role of the unconscious in creativity. In the next section, I will look at Sigmund Freud’s view that “creative success may depend on deep unconscious processes and even defense mechanisms such as sublimation” (Baumeister et al. 2014, 186). I will also look at contemporary literature on unconscious cognition.

### *3.3.2 Freud’s psychodynamic theory*

Freud tried to apply his psychodynamic theory to creativity via psychobiography, which retrospectively analyzed creative works to explain them under the view that “creative thinking was dominated by primary-process thinking – irrational and primitive and associated with the id and the unconscious – rather than by secondary process thinking – logical and realistic and associated with the conscious ego” (Weisberg 1993, 29). Human instincts form the id, a primary process which is only concerned with survival and achieving pleasure. Images are formed in the mind to satisfy instinctual needs, but not every need can be satisfied. Thus, the ego uses the secondary process to “satisfy needs through logical thinking and problem solving” (ibid). The ego and id are in conflict and the ego dominates the id forcing id expressions into the unconscious. Those id expressions manifest when the ego is temporarily weakened, such as while dreaming, day-dreaming or on drugs. Primary process expressions are dependent on unconscious material and appear irrational or bizarre, but there is also an idea that there is an internal conceptual logic to them based on their connection to primitive drives. Besides the uniqueness of the unconscious

conceptual logic and its bizarre expressions, there is a further aid to creativity here, which is that through art, a creative individual can express these repressed instinctual needs.

Psychobiographies were analyses of great works through a Freudian lens, but they have since been debunked. For example, Freud treated Leonardo da Vinci's Mona Lisa's enigmatic smile as da Vinci's repressed "emotions towards his unknown mother" (ibid., 30) in that the particular model aroused those feelings. But more recent studies show that even earlier works of da Vinci's feature women possessing the same enigmatic smile (ibid., 36), suggesting the smile was more a feature of da Vinci's style than anything particular about the model in Mona Lisa. Psychobiographies unsuccessfully attempted to apply Freudian psychoanalysis to art, but Freudian psychology of the role of the unconscious in cognitive acts is still very relevant today.

Analyzing the role of the unconscious in cognitive acts is a question of whether or not people choose, through acts of will, their various actions and behaviors or if these actions and behaviors come from environmental cues or internal unconscious forces. Freud considered actions and behaviors products of unconscious primal urges. This section mainly looked at the history of the passive view, according to which creative acts 'happen' to agents. Even if the agents are consciously aware of their "gift," this does not imply that they deserve credit for their creativity. Today most people would not accept the divine or the hand of nature as explaining the genesis of creative ideas. But creativity could still be understood in somewhat similar terms. For example, Bargh & Chartrand point out that "the mainstream of psychology accepts both the fact of conscious or willed causation of mental and behavioral processes and the fact of automatic or environmentally triggered processes" (Bargh & Chartrand 1999, 463).

### *3.3.3 Bargh & Chartrand's view*

For Bargh and Chartrand, conscious processes are "mental acts of which we are aware, that we intend, that require effort, and that we can control if we choose" (ibid.). Certain mental skills require an act of will to initiate them. The more you develop a skill like improvising jazz, the more automatic that experience becomes when you initiate it into action. This explains why improvisation can appear as 'skilled coping' (or habitual action) and still be hampered under a light cognitive load. The ability of jazz improvisation is partly conscious and partly unconscious ability. Aspects of it remain under conscious control, while other aspects have been handed off to the unconscious.

The interesting thing about Bargh and Chartrand's research is that they show how goals and evaluations can be automatic and even primed. In these cases, the will is bypassed altogether. "The goal, once activated, should operate to produce the same effects as if it had been consciously chosen" (ibid., 469). This goal automaticity relates to previous motivations, which basically means the conscious mind has created the pattern of behavior that the unconscious picks up on and eventually takes over. A similar process happens with evaluations as well.

A substantial body of evidence has now accumulated that one's evaluations often (if not usually) become activated directly, without one needing to think about them, or even be aware that one has just classified the person or event as good or bad (ibid., 474).

Evaluations of events or objects eventually become properties of their perceptual representation. We may be evaluating without consciously intending to evaluate. What is especially relevant for the core sense of creativity is how vital the evaluative component in creativity (judgments of taste) has been thus far in seeking a role for consciousness in creativity. It may be the case, however, that when encountering novel artifacts "what we thinking we are doing while consciously deliberating in actuality has no effect on the outcome of the judgment" (ibid.).

Bargh and Chartrand's research de-emphasizes the role of purposefulness and evaluative ability as conscious elements in creativity. These conditions along with understanding and judgment are necessary conditions that creative actions must exhibit, according to Berys Gaut. In the final section of this chapter, I will look at Gaut's agency view and ask if each condition he gives is necessary for creativity and if consciousness is necessary for each condition.

### **3.4 Berys Gaut's agency conception of creativity by flair**

There may be another way to answer definitively whether or not consciousness is a necessary component of creativity. To do this, I will examine Gaut's agency conception of creativity by "flair", which expands on the standard view. Gaut claims that creativity is a property of agents, and that specific agents' actions count as creative. Gaut introduces the concept of "flair", a term which sums up the conditions he gives for the agent's action that is to count as creative. The kind of actions that are creative are "ones that exhibit at least a relevant purpose, some degree of understanding or skill, a degree of judgment and an evaluative ability directed to the task at hand" (Gaut 2010, 1040).

Therefore, according to Gaut, creativity is “the capacity to produce original and valuable items by flair” (ibid 1041). I will look at each of the conditions and then ask whether or not consciousness is a necessary feature of each one.

### *3.4.1 Creative actions must exhibit purposefulness*

If we accept the minimal accounts of agency, then all creatures act with a purpose (the fundamental purpose ‘to be’). However, is purposefulness in some more demanding sense necessary for creative acts? Purposefulness in creativity relates to a certain intentional state and this presupposes a mental state. As the last section illustrated, not all mental states are conscious, however; so purposefulness may not require consciousness, even if it requires a mental state.

To show the importance of purposefulness, Gaut draws the contrast between a putative creative object made purely by chance versus one made intentionally. He claims that how something is made “plays an essential role in determining whether the act of making it is creative” (Gaut 2003, 150-1). We do not classify something made purely by luck as creative. For example, suppose that I drop my guitar down a flight of steps and on the way down the guitar strikes each step producing a single note each time; when it finally lands, the successive notes are a melody I find beautiful. Have I been creative? Gaut would say that this was pure luck. In order for my creation to count as creative, it must have been my conscious intention to produce a melody.

However, if I flip a paintbrush on a canvas and I like what I see, then although I did not intend exactly what happened, I can still be considered creative. Berys Gaut says that “this follows the reasoning that actions are ‘intentional under some description’” (Gaut 2018, 130)

It is not part of my case that to be creative an artist must intend to produce all of the artistic values that her work exemplifies. That would entail that if an artist were not aware of a particular value of her work, say the intensity of the interaction between two colors in a paintings background, we could not give her creative credit. That is too strong a requirement (Gaut 2018, 130).

Let us grant, then, that purposefulness is necessary for creativity. But is consciousness necessary for purposefulness? In the last section, Bargh and Chartrand pointed to cases where goals could be primed by the environment or even habituated by repeated motivations. In these cases, it would seem that consciousness is not necessary for having a goal despite the initial motivations having been initiated by an act of will.

### *3.4.2 Creative actions exhibit a degree of understanding or skill*

Gaut also says that we must rule out cases of creativity in which the product came about by a mechanical procedure. This is because any creative solution that is realized mechanically shows no degree of understanding or skill. Gaut gives the example of Charles Goodyear's discovery of vulcanization (Gaut 2010, 1041). The theory of vulcanization is considered a highly original and valuable scientific discovery, but Goodyear's discovery is not considered a creative act because of the manner in which he made his discovery. Goodyear apparently dropped substances into liquid rubber until he found something that caused a reaction - sulfur. This example shows that he lacked an understanding of what he was doing, and therefore he was not creative. The discovery is a case of good luck since he simply could have continued dropping substances in liquid rubber until he got a reaction.

Is having an understanding of what you are doing necessary for creativity? In the cases of math and sciences, perhaps, because otherwise one simply relies on luck for their achievements. One could solve a math problem simply guessing, for example. In the arts and other cases, it is harder to say and I think it depends on how strict the degree of understanding is. For example, it is not obvious to me when I sing that I understand exactly what I am doing. Mostly, my singing comes naturally and the aspects which I utilize (opening my mouth and pushing out "sound" are not so clear). It could also be argued that animal behavior, such as the erratic flight patterns of bats, constitute understanding (knowing how avoiding predators), but it is not clear the degree that the understanding takes. It may be a reaction, much like 'skilled coping', which contains no conscious experience. It may be odd to say that the bat "knows" what it is doing, but it seems more odd to say "it has no idea" what to do in relation to predators.

### *3.4.3 Creative actions must exhibit a degree of judgment*

In order to motivate the third "flair" condition, Gaut gives the example of someone who produces a painting simply by following a rule, such as paint-by-numbers (Gaut 2010, 1041). Doing this leaves no room for individual judgment, therefore this action is not creative.

Do creative actions exhibit a degree of judgment? Yes: exercising judgment is necessary for your work to be *your* work, in the relevant sense, rather than somebody

else's work. And the work must be *your* work, in order to be your creative work. The relevant sense of "creativity" includes a certain kind of relationship between the creator and their artifact. Because of these relationships, creators are praised for their creative achievements. As Harry Deutsch says of the paint-by-numbers case "all the creative work has been done for them; and the same is true for the plagiarist" (Deutsch, 1991, 211). However, is consciousness necessary for judgment? No, and this is again for the same reason: that there can be primes or environmental cues which activate your goals and evaluative abilities without you being consciously aware.

*3.4.4 Creative actions must exhibit an evaluative ability directed to the task at hand*  
The example given in support of this condition is of a chimpanzee painting (Gaut 2010, 1041). The chimpanzee adds paint until the trainer removes the painting at a time in which it is aesthetically pleasing. The issue is that the chimpanzee does not show an evaluative ability; the trainer does. Because chimpanzees have no evaluative ability, they will continue to add paint until the canvas is covered. Evaluative ability is essentially 'taste' (what one likes). The chimpanzee's actions are not creative actions; they may be intentional actions, but they do not originate from an intention to create a painting as we understand it (though they originate from the intention to make a mess).

Is exhibiting an evaluative ability directed to the task at hand necessary for creativity? I do think that evaluative ability is a necessary condition of creativity. If one cannot evaluate their work, then creativity has not taken place. Even if the determination of evaluation is that the work is bad, one is making an evaluative judgment on whether or not the work meets their evaluative standard. Without this ability, the actions one is doing do not seem to be creative actions.

Is consciousness necessary for evaluative ability? No, for two reasons. The first is the same reason as the reason in the case of judgment: unconscious primes and environmental cues can trigger 'good' or 'bad' judgments even when one is consciously thinking otherwise (Bargh & Chartrand 1999). The second reason is that a thermostat can be said to have an evaluative ability: it can perceive changes in the temperature of the room and adjust itself accordingly. Similarly, a computer can instantiate style, which is the functional equivalent of an evaluative ability. A style is "a general pattern of ideas/artifacts that is sustained over time by the people adopting it" (Boden 2014,

230). An example would be the sculptor William Latham's genetic algorithm program which generates sustained visual styles (ibid).

This chapter has been analyzing the role of consciousness in creativity. As it stands, I cannot find any necessary role for consciousness in creativity. The praiseworthiness of creativity rules out passive views, according to which creativity is "gifted" by gods, madness or nature. The imagination is not necessary for creativity, as many examples of passive creativity show. I found that both purposefulness and evaluative ability could be primed and triggered by environmental cues, without conscious involvement. The role of taste is not entirely clear. Even the Surrealists, for example, found their "unconsciously generated artworks attracted their conscious approval once they had been formed" (Boden 2014, 235). However, taste is not a productive ability *and* as it was shown a computer can instantiate styles and moods, effectively simulating taste. The next chapter deals with John Searle's position – that computer instantiation of taste is not *really* taste – and looks at some counterarguments to the claim that 'Computers cannot be conscious'.

## 4. COMPUTERS CANNOT BE CONSCIOUS

This chapter will assess the claim that computers cannot be conscious. I have argued above that creativity does not require consciousness, i.e. that the first premise of the master argument against computer creativity is false. If that is indeed the case, then the argument fails regardless of the status of the second premise. However, since the first premise was not quite conclusively refuted, it is worthwhile to take a look at the second premise as well, just in case the first premise can be vindicated, after all.

In this chapter, I will look at John Searle's Chinese room thought experiment and some of the debate surrounding it, such as the debate on whether or not the Turing Test functions as an effective measure of intelligence. Searle was addressing his argument against the claims of classical AI theorists, that an AI that simulates "understanding" would actually understand the way that humans understand. Given Searle's position, a computer simulation of 'taste' is not *really* taste, and this perhaps could mean that a computer cannot be genuinely creative, because it lacks a necessary conscious element that is required for genuine creativity. I will also respond to the assumption that as a programmed artifact, a computer cannot be conscious, since one may use this view to defend the second premise of the master argument. Finally, I will present some objections to the second premise from Daniel Dennett and Hilary Putnam.

### 4.1 Searle and the Chinese room

Using the Chinese room thought experiment, John Searle argues that a machine could think only if it had "internal causal powers equivalent to those of brains" (Searle 1980, 1). Searle is assuming that intentionality, or 'aboutness,' is a product of causal features of the brain. Searle's argument looks like this: computer programs are syntactic formal systems; human minds are semantic (thoughts have meaning); "syntax by itself is neither constitutive of nor sufficient for semantics"; and therefore, "programs are neither constitutive of nor sufficient for human minds" (Churchland and Churchland 1990, 33). The question of going from syntax to semantics is what is at stake in Searle's paper. Computers run programs to simulate thought, but programs are not 'thinking'. They simulate thinking, but that does not constitute real thinking, real understanding.

To illustrate this point, Searle imagines a monolingual English speaker in a room. In the room, and written in English, is a rulebook for outputting a string of Chinese symbols in response to different Chinese symbols. Into the room comes one set of symbols and out goes another. The English speaker is providing adequate responses to the outsiders' questions, but at no point does the English speaker "understand" Chinese. However, for those outside the room, it looks like the operator in the room understands Chinese.

It is obvious, according to Searle, that the man in the Chinese room does not understand Chinese. This is also what both Searle and Paul and Patricia Churchland object to about the Turing Test, in general. The Turing Test is a very demanding behavioral test for intelligence, or thinking, in which a machine engages in conversation with a human and passes the test if it convinces the human they are talking to another person. The Chinese room thought experiment is supposed to provide a counter-example to the Turing test: a system that passes the test, but clearly does not think (because it lacks understanding).

Searle's Chinese room was an attempted refutation of the "strong AI" claims that certain symbol manipulating computer programs could be said to "understand" questions that they gave answers to. Searle had claimed that "syntax by itself is neither constitutive of nor sufficient for semantics" (ibid., 33) but to the Churchland's, this was question-begging similar to the objections that "small particles by themselves could not constitute or be sufficient for the objective phenomenon of light" (ibid., 34) or that inanimate matter could constitute life. They state that Searle's claim that rule-governed symbol manipulation will never make up semantic phenomena is false. The question they deem important to ask regarding conscious computers is: what are the relevant causal powers for conscious intelligence? This is still an open question, but Searle and the Churchland's both agree that the Turing Test is insufficient as a test for conscious intelligence on the grounds that "it is very important how the input-output function is achieved" (ibid., 37). Searle's central point that is relevant for the current purposes is that a simulation of consciousness is not real consciousness. Mistaking a simulation of consciousness for real consciousness is troubling for many reasons. Humans might treat the simulated consciousness as real consciousness, which means caring for it, granting entities that simulate consciousness autonomy, rights or personal freedoms. On the other hand, that might not be a bad thing: treating entities

that simulate consciousness as if they had consciousness might be considered a practical and safe position.

In the next section, I will discuss the Daniel Dennett's view about the Turing Test. In contrast to Searle, he considers the Turing Test valuable. He also warns against the overestimation of cognitive powers in computers, especially as we become more and more reliant on computers for our daily lives.

## **4.2 Daniel Dennett and the Turing Test**

Usually when we want to understand something, we reverse-engineer it. But it seems that this cannot be done with consciousness. Consciousness, as subjective experience, cannot be understood by any reductionist model of the brain; and no analysis of causal states and their function in the brain is complete without the subjective experience. Thomas Nagel says the reason for this “is that every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective, physical theory will abandon that point of view” (Nagel 1974, 437). The difficulty in emulating the brain is not just in understanding its architecture, then. For starters, repurposing artificial materials to do the same causal/functional work that the brain does would “dwarf the entire scientific and engineering resources of the planet for millennia” (Dennett 1994, sect.4). Additionally, no real understanding exists of why/how consciousness emerges in the first place. This is what David Chalmers calls the “hard problem” of consciousness (Chalmers 1995, sect. 2). Chalmers distinguishes the hard problems from the easy problems, with the latter being “reportability of mental states, ability of a system to access its own internal states, integration of information by a cognitive system” (ibid.), and so on. The aspects of consciousness mentioned in connection with the “easy problems” are, for the most part, already instantiated by computers.

The difficulties with ascertaining whether computers can have consciousness make the Turing Test potentially valuable, because the test might be thought to provide a way of determining about a given computer whether it is conscious. In the Turing Test (Turing 1950), a human and a computer sit behind closed doors (much like the Chinese room). On the other side sits a judge who converses with both of them without seeing them. If the judge performs no better than chance in determining which interlocutor is human and which one is the computer, then the computer is said to

have passed the Turing Test. Turing “proposed that any computer that can regularly fool a discerning judge in this game would be intelligent - would be a computer that thinks - beyond any reasonable doubt” (Dennett 1985, 296). If one “can think in one’s own peculiar style well enough to imitate a thinking man or woman, one can think well, indeed” (ibid.)

Dennett says Turing’s test was meant to get philosophers to agree on what thinking was, not necessarily as a perfect means to identify a thinking thing, let alone a conscious thing (there might be inhuman ways of being intelligent). The Turing test, when done correctly, means a computer must employ knowledge about the world, politics, society, etc. just to make sense of a sentence. Rules (syntax) of grammar or vocabulary will not guarantee the right understanding (semantics) of a passage or a question. “What fixes the right reading is knowledge about the world,” says Dennett (ibid., 299).

Dennett concludes that the value of the Turing Test lies in the value of our theories on intelligence. “Verdicts on their intelligence or capacity for thought or consciousness would be only as informative and persuasive as the theories of intelligence or thought or consciousness the verdicts are based on” (Dennett 1985, 310). The Turing Test is not perfect, but how can you test for the thing (consciousness) that is only revealed to the subject experiencing it? According to Dennett, over-valuing the cognitive powers of computers could create real social problems. So we need an appropriately hard test for computers’ cognitive powers; and the Turing Test is a good candidate.

For current purposes, it is important that Dennett’s defense of the Turing Test motivates testing computers similarly for consciousness, not just intelligence/thinking. And although this is a hard test, it is one that computers can potentially pass. So it serves as a criticism of the second premise, that computers cannot be conscious.

### **4.3 Further objections to the second premise**

Some argue that computer consciousness is a contradiction in terms because computers do what they are programmed to do and this is the antithesis of consciousness. In reality, computers implement a wide variety of programs running

'algorithms' ("sequence[s] of instructions telling a computer what to do" [Domingos 2015, 1]) and not all algorithms strictly follow rules. Genetic algorithms, for example "can make random changes in the program's own task-oriented rules" (Boden 2014, 230) by following a path laid out by evolution i.e. interacting genes searching for maximum fitness. The fitness function is key to a genetic algorithm and it operates much like selective breeding where humans play the major role in determining the value of an item. Genetic algorithms also mimic sexual selection by creating parent hierarchies and then crossing over their information to create fitter generations (or fitter programs) (Domingos 2015, ch. 5). This elaborate process makes it impossible for the programmer to predict just what a computer will do. So the argument that computers cannot be conscious, since as programmed entities they just follow the rules put into them by the programmer, fails: programmed entities do not necessarily follow the rules envisioned by somebody else.

In this last section of the chapter, I will consider some further objections to the claim that computers cannot be conscious. To start with, Dennett (1994) addresses a couple of reasons someone might give for why a conscious computer is impossible. The dualist position states simply that artificial beings like computers are material things and consciousness is immaterial mind-stuff. This idea borders on mysticism, according to Dennett. Thales, one of the pre-Socratics, thought the loadstone had a soul, and in his time that was the best proto-scientific understanding of magnetism. Consciousness is mysterious now, just like magnetism was in the past, but that does not mean there is not a natural physical explanation for it. History is filled with mysterious or miraculous phenomena which eventually became understood by science.

A second way one might object to the possibility of conscious computers has to do with the vitalist claim that consciousness is dependent on organic material. Since computers are inorganic, they cannot be conscious. This claim is species-centric and does not allow for inhuman or alien kinds of consciousness. It is most likely false (because it gives too narrow a conception of consciousness), but can still be defended on the grounds that to mechanistically reproduce all the powers of organic compounds in an alternate medium would require scales that border on the impossible. But even the complexity issue does not suffice to render the original claim unfalsifiable, for

there are plenty of complicated parts that can easily be substituted for simpler models, e.g. artificial heart valves.

Hilary Putnam looks at similar arguments that all seem to share some baseline assumptions about computers, namely that they (computers, though he uses ‘robots’) are programmed artifacts, and anything that is a programmed artifact cannot be conscious (Putnam 1964, 680). I introduced this claim in the beginning of the section by looking at how computer programs presently function and how their “programming” does not constitute predictability. Interestingly, Putnam imagines an alien race appearing and claiming that we (human beings) were programmed artifacts and they (because they created us) know our every thought. Using the above logic, it should follow that since we are now programmed artifacts, that we were not conscious.

Putnam also considers what he calls ‘Ziff’s argument’ (ibid., 687). In Ziff’s argument, we come across a plant and, upon closer inspection, it turns out to be composed of tubes, gears and wheels. It is more a mechanism than a plant. Now, given how we understand the meaning of the word ‘alive’, we can say “anything that is not alive is not conscious.” According to Ziff, we also tell whether or not something is alive based on its structure, not just its behavior. So the plant-like mechanism composed of tubes and gears is not alive, under this definition, therefore it cannot be conscious. But this understanding of ‘alive’ is too dogmatic, according to Putnam, and bases the idea of a “living thing” on privileging “soft stuff” over “hardware” (ibid, 686).

This chapter looked at John Searle’s Chinese room experiment which tried to refute the claims of strong AI theorists that a symbol manipulating machine ‘understood’ just like a human understands. The Churchlands and Searle both reject the Turing Test as sufficient for measuring consciousness and related phenomena. Daniel Dennett argues for its validity on the grounds that no other test has yet been imagined which would constitute a comprehensive enough show of intelligence to justify itself the way the Turing Test does. I also looked at some further objections from Dennett and Hilary Putnam to the claim that computers cannot be conscious.

## 5. CONCLUSION

The question of whether or not computers can be creative is still open. I focused only on one argument against computer creativity: (1) creativity requires consciousness; (2) computers cannot be conscious; therefore, (3) computers cannot be creative. To better understand “creativity”, I looked at a core, intuitive sense of “creativity”, one which seems to imply that creativity must be the product of an agent’s actions since we justifiably give creators credit for their creative achievements. Though creativity consists of three parts (agent, process and artifact), understanding creativity particularly requires understanding of how the agent realized their creative ideas.

I further focused on the first premise: ‘Creativity requires consciousness’. Historically, traditional views of creativity have given credit to the gods or divine madness, while the agent was the lucky beneficiary of the “gift”. While this view does not rule out a necessary role for consciousness in creativity, it downplays the agent’s role to one of recipient. Kant’s “genius view” also granted that the creative agent could be gifted with natural ability, but added that a creative agent must have taste, to distinguish their productive works from gifted nonsense.

Taste, however, seems to imply conscious experience, that there must be something it is like to make evaluative judgments. I argued against this supposition about taste. A thermostat can be said to instantiate an evaluative ability, that of perceiving changes in temperature and implementing a programmed response. Additionally, I found that both purposefulness and evaluative ability could be primed and triggered by environmental cues. It is not entirely clear what the role of consciousness in taste is. The Surrealists, for example, found their “unconsciously generated artworks attracted their conscious approval once they had been formed” (Boden 2014, 235). However, taste is not a productive ability *and* as it was shown a computer can instantiate styles and moods, effectively simulating taste.

I put less emphasis on the question of whether computers can be conscious. I looked at Searle’s Chinese room experiment which was given to motivate the claim that simulation of a mental state does not qualify as the real thing. For example, simulated taste is not *real* taste. However, the main claim of Searle’s argument rests on the axiom that no amount of syntax can generate semantics. The Churchlands reject

that claim on the grounds that many naturally occurring phenomena operate similarly, such as light particles illuminating a room or inanimate matter constituting life. For them, the relevant mental processes are important. This is why they both reject the Turing Test as a test for intelligence or consciousness. However, Daniel Dennett shows that the Turing Test is not meant to prove intelligence or consciousness, but merely to serve as defeasible evidence for intelligence or consciousness. Given the subjective nature of consciousness, there may be no effective test to measure it completely reliably. But the Turing Test is robust enough that should a computer pass it, it may justify us thinking of them as conscious.

In conclusion, I argued against the view that consciousness is necessary for creativity. However, if it should turn out that consciousness is necessary for creativity, after all, then (I argued) the possibility of computer consciousness can be defended.

## REFERENCES

- Bargh, John A. & Chartrand, Tanya L. (1999). "The unbearable automaticity of being." *American Psychologist*. Vol. 54, no. 7, pp. 462-479.
- Baumeister, Roy F., Schmeichel, Brandon J., Dewall, Nathan C. (2014). Creativity and consciousness: evidence from psychology experiments, in Elliot Samuel Paul, and Scott Barry Kaufman (eds) *The Philosophy of Creativity: New Essays*, Oxford University Press, New York, NY, pp. 185-198.
- Bharara, Preet. "Putin, Pawns and Propaganda (with Garry Kasparov)". Podcast. Stay Tuned with Preet. CAFE, WNYC Studies, Pineapple Street Media. New York. December 7th, 2017.
- Blackburn, Simon. (2014). "Creativity and not-so-dumb luck". Paul, E.S. & Kaufman, S.B. (eds.), *The Philosophy of Creativity: New Essays*, Oxford University Press, New York, NY, pp. 147-156.
- Boden, Margaret. (2016). *AI: Its Nature and Future*. Oxford University Press. Oxford, UK.
- Boden, Margaret. (2010). *Creativity and Art: Three Roads to Surprise*. Oxford University Press. Oxford, UK.
- Boden, Margaret A. (1998). "Creativity and artificial intelligence". *Artificial Intelligence* 103. pp. 347-356.
- Boden, Margaret. (2014). "Creativity and artificial intelligence". Kaufman, S.B. & Paul, E.S. (eds.) *The Philosophy of Creativity*. Oxford University Press, New York, NY. pp. 224-244
- Bostrom, Nick. (2014). *Superintelligence*. Oxford, Oxford University Press.
- Bukowski, Charles. (2003) sifting through the madness for the Word, the line, the way. Harper Collins.
- Burnham, Douglas. Immanuel Kant: Aesthetics. Internet Encyclopedia of Philosophy. Accessed on March 30, 2018. <https://www.iep.utm.edu/kantaest/#SH2d>
- Carruthers, Peter. (1989). "Brute Experience," *The Journal of Philosophy*, LXXXVI, 5 (May 1989) pp. 505-516.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge University Press.
- Chalmers, David. (1995). "Facing up to the problem of consciousness". *Journal of Consciousness Studies* (2)3, pp. 200-219. Accessed on March 28, 2018.

- Chalmers, David. (2010). "The singularity: A philosophical analysis". *Journal of Consciousness Studies* 17: (pp. 7-65).
- Chan, Dawn. (2017). The AI that has nothing to learn from humans. Atlantic.com. October 20, 2017. Accessed on March 10, 2018. <https://tinyurl.com/y7ucmuzo>
- Cohen, Ted. (2003). "The inexplicable: Some thoughts after Kant." Gaut, B. & Livingston, P., eds., *The Creation of Art*. Cambridge University Press, New York. pp. 138-147.
- "Computer". Etymonline.com The Online Etymology Dictionary. Accessed April 12, 2018 from <https://tinyurl.com/yc34pttt>
- Dennett, Daniel. (1985). "Can Machines Think?" Shafto, M (ed.) *How we know: nobel conference XX*. Harper & Row, San Francisco, CA 1985.
- Dennett, Daniel. (1994). *Consciousness in human and robot minds*. Ito, et al., eds., *Cognition, Computation and Consciousness*, downloaded from: <https://tinyurl.com/y7s48r8l>
- Domingos, Pedro. (2013). *The Master Algorithm*. Basic Books. New York, NY.
- Deutsch, Harry. (1991). "The creation problem". *Topoi* 10. pp. 209-225
- Gaut, Berys. (2003). Creativity and Imagination. In B. Gaut, & P. Livingston (Eds.), *The Creation of Art: New Essays in Philosophical Aesthetics* (pp. 148-173). Cambridge University Press.
- Gaut, Berys. (2010). The philosophy of creativity. *Philosophy Compass*. May 12, 2010, (pp. 1034-1046).
- Gaut, Berys. (2018). "The value of creativity". *Creativity and Philosophy*. Gaut, B. & Kieran, M. (eds.) Routledge. New York.
- Hills, Alison & Bird, Alexander. (2018). "Creativity without value". *Creativity and Philosophy*. Gaut, B. & Kieran, M. (eds.) Routledge. New York.
- Hoefer, Carl, "Causal Determinism", *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), <https://tinyurl.com/yd6f68zl>
- Kasparov, Garry. (2017). *Deep Thinking*. PublicAffairs. New York.
- Kant, Immanuel. (1987). *Kant's Critique of Judgement*. Trans. by J.H. Bernard. (2nd ed. revised) London: Macmillan, 1914.
- Kaufman, B. & Paul, S. (2014). "Introducing the philosophy of creativity". *The Philosophy of Creativity*. Paul, E.S., and Kaufman, S.B. (eds.) Oxford University Press. New York, NY. pp. 3-14.

- Kieran, Matthew. (2014). "Creativity as a virtue of character". *The Philosophy of Creativity*. Paul, E.S., and Kaufman, S.B. (eds.). Oxford University Press. New York, NY. pp. 125-143
- Kronfeldner, Maria. (2009). Creativity naturalized. *The Philosophical Quarterly*, Vol 59, No. 237, pp. 577-598.
- Kronfeldner, Maria E. (2014). *Darwinian Creativity and Memetics*. London: Routledge, 2014. eBook Collection (EBSCOhost), EBSCOhost (accessed November 13, 2017).
- Metz, Cade. (2016). "Why the final game between Lee Sedol and AlphaGo is such a big deal for humanity". Wired.com. March 14. <https://tinyurl.com/y82vdl93>
- Moravec, Hans. (1988). *Mind Children*. Harvard University Press. Cambridge, MA.
- Nagel, Thomas. (1974). "What it is like to be a bat?" *Philosophical Review* 83: pp. 435-456.
- Nanay, Bence. (2014). "An experiential account of creativity", in Elliot Samuel Paul, and Scott Barry Kaufman (eds), *The Philosophy of Creativity: New Essays*, Oxford University Press, New York, NY, pp. 17-35.
- Nietzsche, Friedrich. (1878). *Human All Too Human*, trans. by Collingwood, R.J.
- Poe, Edgar Allan. (1850). "The Poetic Principle" (reprint), *Home Journal*, no. 36. August 31. Pp. 1-6
- Popova, Maria. (2012). "How Creativity Works, From Nietzsche to Jonah Lehrer". *Atlantic.com*. 3.22.12. Accessed: March 14, 2018. <https://tinyurl.com/ybyq7sl9>
- Putnam, Hilary. (1964). "Robots: machines or artificially created life?". *The Journal of Philosophy*, Vol. 61, No. 21, American Philosophical Association. November 12, 1964. pp. 668-691.
- Rothenberg, Albert & Hausman, Carl R. (1976). *The Creativity Question*. Duke University Press, Durham, NC, USA.
- Schlosser, Markus. (2015). "Agency", *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/fall2015/entries/agency/>
- Turing, A.M. (1950). "Computing machinery and intelligence", *Mind*, Volume LIX, Issue 236, 1 October 1950, Pages 433–460
- Weisberg, Robert W. (1993). *Creativity: Beyond the Myth of Genius*. W.H Freeman and Company. New York.

Wright, Robert. (1996). "Can Machines Think?". Time.com. March 25, 1996. Accessed March 10th, 2018. <https://tinyurl.com/y88k5m5j>

## **ABSTRACT**

### **CAN COMPUTERS BE CREATIVE?**

[Kas arvutid saavad olla loovad?]

It is sometimes claimed that computer automation of work will free people up to be more creative and that being creative is a way to add value to one's life. However, if computers themselves become creative, then this might impede a human's ability to add value to their life. This thesis investigates the question 'Can computers be creative?' by assessing the role that consciousness plays in creativity. Specifically, I assess the following argument against computer creativity: (1) creativity requires consciousness; (2) computers cannot be conscious; (3) therefore, computers cannot be creative. The first premise is challenged by cases where humans program computers to produce apparently creative (novel and valuable) results. However, in the relevant sense of "creativity", creativity must result from the actions of an agent. Accordingly, I will focus on whether the agency that is required for creativity requires consciousness. I draw on literature from philosophy of creativity, philosophy of artificial intelligence and philosophy of mind to argue that consciousness is not necessary for creativity. I also present some considerations against the second premise, that computers cannot be conscious.

## Lihlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Patrick Keenan,  
(autori nimi)

1. annan Tartu Ülikoolile tasuta loa (lihlitsentsi) enda loodud teose

'Can Computers be Creative?'  
(lõputöö pealkiri)

mille juhendaja on Eve Kitsik,  
(juhendaja nimi)

- 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus/Tallinnas/Narvas/Pärnus/Viljandis, **pp.kk.aaaa**

**Non-exclusive licence to reproduce thesis and make thesis public**

I, Patrick Keenan,  
(author's name)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:
  - 1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and
  - 1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

'Can computers be creative?'  
(title of thesis)

supervised by Eve Kitsik,  
(supervisor's name)

2. I am aware of the fact that the author retains these rights.
3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu/Tallinn/Narva/Pärnu/Viljandi, **dd.mm.yyyy**