THE AESTHETIC VALUE OF ARTIFICIALLY GENERATED ART

CARLIJN BUIS

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE IN COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE

DEPARTMENT OF
COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

STUDENT NUMBER

2055420

COMMITTEE

prof. dr. Max Louwerse
Eriko Fukuda

LOCATION

Tilburg University
School of Humanities and Digital Sciences
Department of Cognitive Science &
Artificial Intelligence
Tilburg, The Netherlands

DATE

May 19, 2023

WORD COUNT

6196

ACKNOWLEDGMENTS

I want to thank my supervisor dr. prof. Max Louwerse for his advice and assistance during the process of writing my thesis. I sincerely appreciate the time and effort he invested in providing constructive feedback and offering valuable suggestions

THE AESTHETIC VALUE OF ARTIFICIALLY GENERATED ART

CARLIJN BUIS

Abstract

This study examines the aesthetic evaluation of artwork created by humans as opposed to artwork created by artificial intelligence, along with the effect of the Rule of Thirds on aesthetic perception. The experiment included carefully selected paintings made by humans and paintings generated with AI and these materials were further divided into those that followed the Rule-of-Thirds and those that did not. In the experiment, subjects evaluated these paintings on aesthetic value. The results concluded that paintings created by human artists have a higher level of aesthetic appeal than those produced by AI algorithms. In addition, paintings that followed the Rule-of-Thirds were consistently evaluated to be more aesthetic than those that did not. This study adds to the discussion on the value of AI-generated art and reinforces the Rule-of-Thirds' positive effect on aesthetic evaluation.

1 DATA SOURCE, ETHICS, CODE, AND TECHNOLOGY STATEMENT

The data collected in this research has been acquired from human participants through Qualtrics (Qualtrics, Provo, UT) via an online request. The data was collected especially for this experiment. The obtained data is anonymised. Half the materials for the experiment were generated using Dall·E 2. The other materials were paintings collected on the internet from two human artists, and those are all images in the public domain. The code was written by the author. The coding application used to generate results was R Core Team (2020). In this application, the following packages were used: tidyverse (Wickham, 2020), stats (R Core Team, 2020), coin (Hothorn et al., 2021), and dplyr (Wickham et al., 2021). I have also used Overleaf as a latex operator.

2 INTRODUCTION

Using AI tools and techniques to produce original and innovative works of art has become more and more common for artists (Chatterjee, 2022). With the newest Artificial Intelligence (AI) developments, it is possible to use AI to make images, music, and movies (CineD, 2023; OpenAI, 2022; Tan & Li, 2021). These developments open up a whole new field of possibilities, blurring the line between human and machine output. The field of AI art is fast developing, and it is still unclear how it will affect both the art world and society at large (Epstein et al., 2020; Eshraghian, 2020). It has therefore become imperative to ask ourselves what we as humans value in art. This thesis contributes to this ongoing discussion and provides new insights into the value of artificially generated art.

Since the development of generative adversarial networks (GANs), (Goodfellow et al., 2014), the use of AI for the creation of art has skyrocketed. Dall·E 2 (OpenAI, 2022) is the latest development in this area and has proven very effective at creating art using natural language prompts. Before these developments, the creation of art has been characterised as a uniquely human ability (Zaidel, 2010). However, with Dall·E 2, anyone can create unique art pieces in a matter of seconds. Consequently, the art world is likely to be affected by the emergence of these AI techniques (Epstein et al., 2020; Eshraghian, 2020).

In order to look into the value of art, it is vital to examine the concept of aesthetics. Aesthetic value can best be described as the quality of an art piece that is based on the ability to evoke a pleasing sensory or emotional experience (Feagin & Maynard, 1997). The Aesthetic value of an art piece is determined by a multitude of factors, including the spatial composition (Obrador et al., 2010). A common technique for creating a harmonious spatial composition is called the Rule-of-Thirds (RoT) (Koliska & Oh, 2023). The RoT states that when you divide an image by nine equal rectangles, using two horizontal and two vertical lines and then placing a visually important element at the intersection or along the gridlines, will create a more harmonious composition (Koliska & Oh, 2023).

This thesis deals with a topic of importance for a number of reasons. First of all, as aforementioned, recent AI applications have shown tremendous process in the field of creative design. There is growing concern that as AI algorithms and models progress, they may eventually completely replace human creativity (Eshraghian, 2020).

Second, the idea of aesthetics and art is important to human culture and has long been a focus of philosophical study (Feagin & Maynard, 1997). The comparison of the aesthetic value of art produced by humans and by

AI has consequences for our understanding of human creativity and the nature of art, as well as for the art world.

Thirdly, since the art world is likely to be financially affected by the emergence of these AI techniques, the rise of AI-generated art may affect artists financially and therefore have an impact on the art market as a whole (Epstein et al., 2020; Eshraghian, 2020).

With the rapid emergence of AI developments in the field of art, a discussion has sprouted on the value humans put on AI-generated art. Questions are raised on whether these artificially generated art pieces are of the same quality as those created by humans. To gain a better perspective on the discussion, this thesis will compare the aesthetic value of art created by AI and humans when it is unbeknownst to the participants that they are looking at artificially generated art. The research question this thesis will try to answer is:

> *What aesthetic value do humans attribute to art generated by AI versus art created by humans?*

This research also includes the RoT in its experimentation. To get a clearer view on the role of RoT in the evaluation of aesthetics, this thesis includes the sub-question:

> *How does the Rule-of-Thirds influence the evaluation of aesthetics*

As for the first research question, current literature on the aesthetic value given to artificially generated art by humans portrays a bias against artificially generated art versus art created by humans (Chamberlain et al., 2018; Ragot et al., 2020; refer to Related Works section 3.2). Based on the literature, this thesis will explore whether this relatively negative aesthetic evaluation of AI-generated art will remain when the artist is unknown to the subject. Therefore, this thesis proposes hypothesis 1.

H1: *Humans give more aesthetic value to art created by humans, as opposed to art generated by AI.*

When it comes to the assessment of the RoT in terms of its aesthetic appeal, the literature is divided on whether the RoT has an effect on the aesthetic value (Amirshahi et al., 2014; Koliska and Oh, 2023; refer to Related Works section 3.3). For this reason, this thesis proposes hypothesis 2 based on the literature.

H2: *Humans give more aesthetic value to art that follows the RoT.*

The findings indicate that hypothesis 1 on the value of AI-generated art is unsupported. The opposite effect was true, which is that the paintings made by humans were evaluated as more aesthetically pleasing. Additionally, the results support the second hypothesis which states that the RoT increases the aesthetic evaluation of a painting.

3 RELATED WORK

This study focuses on the interaction of artificial intelligence and creativity, with a particular interest in the aesthetic worth of art produced by humans as opposed to AI. The question at hand is whether AI-generated art can equal or exceed the aesthetic worth of art made by humans.

3.1 *Recent AI Advancements*

Firstly, this section looks at the research behind the latest advancements in AI and computer graphics in the area of art. A study by Gatys et al. (2016) has shown that using CNN (convoluting Neural Networks) can be used for generating images. NST (Neural Style Transfer) uses CNN to separate the 'content' and 'style' from an image and to recombine them to produce a new image. This generated image is heavily based on the input sample and will not create truly novel art pieces.

In 2014, Goodfellow et al. (2014) presented the GANs. These GANs train two models simultaneously. One model is generative and captures the input distribution of images and generates new images. The second model is discriminative and is trained to distinguish the 'real images' in the sample distribution and the images that the generative model has created. The goal is to use minimax optimization, so the discriminative model makes as many mistakes as possible, to thus improve the generative ability of the generative model. GANs were able to produce realistic-looking fake art images. However, just like NST, these generated images are not yet able to create truly novel art pieces.

Elgammal (2019) wanted to create an algorithm that was able to create truly novel art pieces and introduced AICAN. Elgammal (2019) achieved this by modifying GANs. They altered the GANs by maximizing the deviation from 'style' and minimizing the deviation from the sample group. Elgammal et al. (2017) states that participants were not able to distinguish the real images from those that were generated.

One of the last major advancements in this area was from Radford et al. (2021) who created the CLIP model. This model was trained using 400 million text-image pairs from the internet. The model uses contrastive learning to learn which text-image pairs are similar and which are not. CLIP then measures the similarity between the text and the image.

In 2021, Dall·E became open-access by OpenAI to selected individuals. A year later in September 2022, Dall·E 2 became accessible to anyone (OpenAI, 2022). These networks can generate images from natural language. The models use the CLIP model to measure and rate how well the generated images match the text. Dall·E 2 is now able to create novel digital art pieces with any given prompt. Dall·E 2 is a powerful AI tool, that will be utilized in this thesis. Part of the paintings from the materials in the experiment have been generated using Dall·E 2.

3.2 *Value of AI generated Art*

There is a connection between art and its artist. It is well known that an artist's status influences the price of their artwork. An issue arises now that AI has entered the art world as an artist, namely the issue of what value could be assigned to artificially generated art. Thus far, two artificially generated artworks have entered the art world and they have exceeded expectations. The first was Edmond de Belamy, a painting printed on canvas that was generated using GANs ("Portrait by AI", 2018). This art piece was the first artificially generated art piece to be sold at an auction. Though the piece was estimated at around $10.000, the piece sold at $432.500. The second piece of art that caused controversy is called PSEUDOMNESIA: The Electrician (Glynn, 2023). This photograph was entered for the Sony World Photography Awards. The photograph won in the creative photo category. However, Boris Eldagsen generated this photograph using Dall·E 2, to test whether these creative art contests were ready for AI-generated art.

These events raise the question of what the general consensus is on the value of AI-generated art. Hannah Kaube and Rahman (2023) have performed a study that investigated the link between art and artist. They found that the paintings of artists with a negative-social biography were received as less aesthetically pleasing, than paintings of artists with a neutral biography. This study demonstrates that artists can influence the aesthetic perception of a painting. This becomes more evident in the study by Ragot et al. (2020). This between-subjects study asked participants to evaluate a painting. Half of the participants were told that the art was generated by AI and the other half was told that the art was made by humans. The results showed that participants evaluated the art labeled as human-made higher than the paintings labeled as AI-generated. These results imply a negative bias towards the paintings generated by AI.

There are several assumptions on why this bias exists against AI-generated art. One might be the aspect of effort and time that went into an art piece. Kruger et al. (2004) investigated the role of effort in the appreciation of art. They indeed concluded that people value the skill, effort, and intention that went into a work of art. This could mean that the bias towards AI is due to AI-generated art being perceived as being made with less effort. Chamberlain et al. (2018) conducted two experiments on this matter. The first also showed a negative bias toward AI-generated art. However, in the second experiment, instead of using AI-generated art, they used a robot. This experiment was split into two settings, In the first setting the participants sat across the robot, while the robot took a picture of the participant and used a robotic arm to create a sketch based on the picture. The second setting did not include this interaction with the robot. Instead, participants were shown pictures drawn by the robot and they were either informed that the robot made these pictures or they were given no source information. The results show that the art produced in setting 1 had a greater positive impact on the aesthetic evaluation of the artwork than in setting 2. This implies that when participants were given the chance to interact with the robot, as is the case in setting 1, then the perceived aesthetic status of the artwork increased. Therefore, this interaction effectively negated the bias of artificially generated art.

In conclusion, the value of AI-generated art is a controversial topic that has raised a lot of discussions. A lot of research has been done on the bias of AI-generated art. However, this leaves the question of whether AI-generated art is still perceived to be less aesthetically appealing when the artist is completely unknown.

3.3 *Rule-of-Thirds*

Aesthetics, as a concept, studies the ideas of aesthetic experience, aesthetic judgment, and the elements that make objects, artworks, or experiences in general appealing, harmonious, and emotionally resonant (Feagin & Maynard, 1997). There are various methods to differentiate between the level of aesthetics in artworks. One way is to look at the spatial composition of a painting. One frequently used technique for the evaluation of spatial aesthetics is called the Rule-of-Thirds (RoT) (Koliska & Oh, 2023). This composition technique can aid to create more visually aesthetic images. The RoT states that when you divide an image by nine equal rectangles, using two horizontal and two vertical lines and then placing a visually important element at the intersection or along the gridlines, will create a more harmonious composition (Koliska & Oh, 2023). This technique is popular in photography and painting. Many artists and experts claim that the Rule-of-Thirds (RoT) is very important in the spatial composition of a photograph or painting (Gooch et al., 2001; Mai et al., 2011), but the RoT might not have the significant effect on the aesthetic perception that was expected (Amirshahi et al., 2014). Research into the role of the RoT towards the subjective aesthetic value of an artwork is still limited.

Koliska and Oh (2023) state that the application of the RoT can help guide the viewer's focus to the image's main element. They concluded that the perceived valence of an image became significantly higher for images that followed the RoT principle. This means that the participant was able to recognise the most integral part of an image more accurately. Another study by Tang et al. (2020) developed a computational method based on the RoT that could asses the composition harmony. They concluded that there was a positive relationship between the method and the compositional harmony. To conclude, the effect of the RoT on aesthetic value is still debated. The experiment in this study will provide additional information about the RoT concerning aesthetics using subjective experiences.

4 METHOD

4.1 *Materials*

The materials for the experiment included 40 paintings (refer to the appendix). This number was partially based on the study by Koliska and Oh (2023), who used a similar experimental design and they decided on 24 images in total. Half of the paintings were created using Dall·E 2. The other half was collected from the internet by two human artists. The two human artists chosen for the materials were Theodore Robinson and Armand Guillaumin. They were chosen for two reasons. First of all, they are relatively unknown, so there is a smaller chance for participants to recognise the painters. This could be problematic, because if they recognise the paintings from the painters, then they might show bias in their answers. Secondly, both painters have enough paintings that fit the required criteria. These criteria help to filter the paintings for the purpose of homogeneity. These criteria are based on the art style, medium, colour palette, and subject matter. the first criterion is that the paintings in the database are from the art movement Impressionism. This movement started in the 19th century. The art style is characterised by its broken brush strokes and the emphasis on the effects of light (Samu, 2004). The medium of these paintings is an oil painting. In addition, The paintings visualise a natural landscape with the inclusion of one or more trees. Lastly, all the paintings are at least one-third green.

Ten paintings were selected from both painters for the materials. The other 20 paintings were generated using Dall·E 2, with the following prompt: "An impressionist oil painting of a landscape with trees". Each iteration of Dall·E 2 with this prompt created four paintings. One painting was selected from each iteration, that best fit the criteria.

After the selection process, the paintings underwent a second selection process. Half the paintings in the database were selected to fit the RoT criteria, while the other half did not meet this criterion. This means that half of the paintings from Dall·E 2 fit the RoT criteria and half the painting from each human painter matched the RoT criteria. The paintings were manually cropped to achieve the desired RoT result. All paintings were then resized to 360 x 280 pixels, to create a homogeneous effect in quality as well. Figures 1 to 4 represent a painting from each of the four groups. These four groups are human x RoT (figure 1), human x no RoT (figure 2), AI x RoT (figure 3), and AI x no RoT (figure 4).

4.2 *Experimental Design*

The experiment was designed in Qualtrics (Qualtrics, Provo, UT) and was conducted in an online environment. During the experiment, the participants evaluated 40 paintings, one at a time. They evaluated the painting based on its aesthetic appeal. The question they saw with every painting was: "Do you find this painting aesthetically pleasing?" The participant could answer that question with: yes or no. They were also encouraged to pick a reply as quickly as possible to capture the initial reaction. This was done by including a timer that would automatically skip to the next painting after five seconds if the participant had not yet decided on a painting. The timer was not visible to the participants. To help the participants get familiar with the layout, a dummy question was provided. At the end of the experiment, there were two control questions. The first asked if they knew the painters Theodore Robinson or Armand Guillaumin, the two painters of the human-made paintings in the materials. The second asked the participants what they believed this experiment was about. This second question was included to find any participants who knew that the experiment revolved around AI-generated paintings. The experiment also recorded the response times.

The experiment is a two-by-two factorial design. The first independent variable is the creator of the artwork, with the levels: AI and human. The second independent variable is the RoT, with the levels: RoT or no RoT. The responses from the participants concerning the aesthetic value (yes or no) and the response times are the two dependent variables.

Figure 1: 'Olive Grove, Capri' by Robinson (1890). The white grid in the painting demonstrates that this painting follows the RoT, as the main focal point is the tree on the right and this tree follows the right vertical grid line.

Figure 2: 'Gust of Wind, le Brusc' by Guillaumin (1911). The white grid in the painting shows how this painting does not follow the RoT, because the focal point, the tree, is in the middle of the painting and does not align with any gridlines or points.

Figure 3: This painting was generated using Dall·E 2 (OpenAI, 2022), with the following prompt: "An impressionist oil painting of a landscape with trees". The white grid in the painting demonstrates that this painting follows the RoT, as the tree, which is the main focal point, follows the right vertical gridline.

Figure 4: This painting was generated using Dall·E 2 (OpenAI, 2022), with the following prompt: "An impressionist oil painting of a landscape with trees". The white grid in the painting shows that this painting does not follow the RoT, because the tree, the main focal point, does not follow the gridlines or grid points.

4.3 *Participants and Data Pre-processing*

79 participants took part in this study. Several actions were taken in order to prepare and process the data for analysis. First, the dataset was removed of any participants who did not finish the experiment, because incomplete data could introduce bias into the dataset. Those who did not give consent for the use of their data were also left out. This left a dataset of 67 participants. This dataset consists of mostly female subjects (Female = 46, male = 20, I don't know = 1). Additionally, the subjects were predominantly highly educated (94% of participants with hbo, Bachelor WO, Master WO, or pHD). The dataset also included missing data. These missing data values were eliminated, in order to maintain the quality and accuracy of the statistical analysis.

The timestamp of each participant's first click was saved in order to examine response times. Any response time under 0.2 seconds was excluded for reason of quality control. Any participants would need at minimum more than 0.2 seconds to evaluate a painting, any less than that might indicate that the participant was rushing through the experiment.

Overall, taking these actions was crucial to guarantee the accuracy and reliability of the data used in this investigation. This data cleansing and transformation enabled a more precise and insightful analysis.

4.4 *Data Analysis*

All statistical analyses were performed using R Statistical Software (R Core Team, 2020). According to the Shapiro-Wilk test implemented in the R stats package (R Core Team, 2020), the response times data did not follow a normal distribution ($p < 0.001$). From various transformations, it was concluded that transforming the data did not have the desired results as the data continued not to follow the normal distribution. For this reason, it was decided that statistical analysis on the response times would use non-parametric tests, namely the Wilcoxon signed-rank test in the R stats package (R Core Team, 2020). This test deals with paired data. This made it suitable for the dataset because each independent variable in the dataset includes measurements taken on a matched pair. In the case of the independent variable Artist, this matched set would be human and AI. It was also tested if the data met all the assumptions of the Wilcoxon signed-rank test. With a boxplot, it was concluded that the data included a lot of outliers. Outliers could increase the variability and decrease the effect size, so the IQR method was applied to detect the outliers. Using this, any data points that were (1,5 * IQR) distance from the median, were removed. These outliers included 115 out of the 1998 total data points.

For the accuracy, binary responses were recorded (yes and no). A test that would suit the binary nature of the dataset would be a logistic regression from the R stats package (R Core Team, 2020). The dataset met the assumptions for the use of logistic regression. To visualise the results in a bar chart, the ggplot2 function was used from the R tidyverse package (Wickham, 2020).

Additionally, there were participants that either recognised a painter (N = 7) or had some idea of what the experiment was really about (N = 8). A separate test was run on the data without these participants and the results did not change the conclusions drawn from the initial analysis. Please refer to the Results section 5.1 for more information on this.

Lastly, two more packages were used in the analysis. The first is the Dplyr package (Wickham et al., 2021), which was used in the pre-processing part of the code. The second is the coin package (Hothorn et al., 2021) that was used to find the effect sizes from the Wilcoxon test.

5 RESULTS

5.1 *Accuracy*

A logistic regression model performed analysis on the dataset to examine the relationship between the accuracy and two predictor variables artist and composition. The sample consisted of N = 67 participants. Figure 5 displays the probability of the accuracy of the two independent variables. The figure shows that the accuracy of AI as an artist is lower than the accuracy of a human as an artist. This implies that the participants evaluated the painting by humans overall as more aesthetic as opposed to the paintings made by AI. Figure 5 also shows the accuracy of the inclusion of the RoT or not. The accuracy appears more divided amongst the RoT variable. This indicates that the paintings that included the RoT were perceived as more aesthetic when the paintings were created by AI, but the opposite appears to be true when the paintings were created by humans.

The independent variable artist was coded: human = 1, AI = 0, and the independent variable composition was coded: RoT = 1, no RoT = 0. The dependent variable was coded: yes = 1, no = 0. The logistic regression model showed that controlling for the subjects, both the artist variable (B = 0.71, SE = .08, p < .001) and composition variable (B = 0.24, SE = .08, p = .003) were significant in predicting of the log-odds of the accuracy. The odds of the accuracy increased 2.03 times for a one-unit increase in the artist variable, while controlling for subjects, 95% CI [1.73, 2.37]. As the coefficient is positive and the log odds are more than 1, this indicates that subjects gave a higher accuracy to human-made paintings. This means

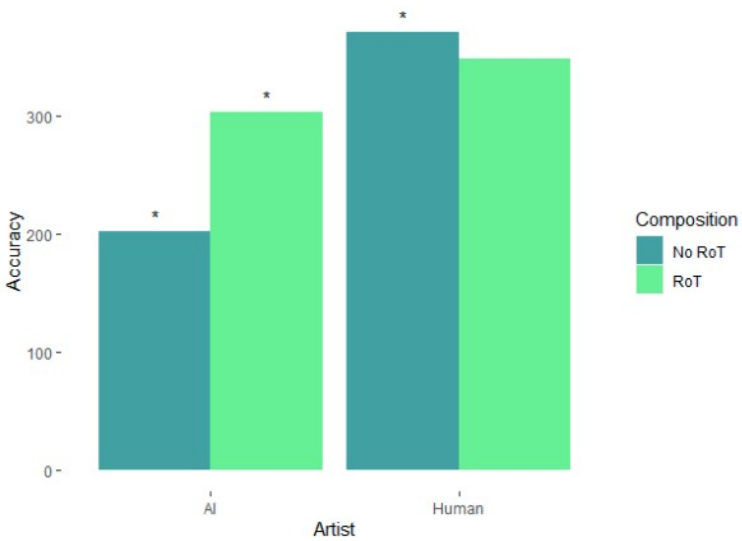


Figure 5: Grouped barchart of the accuracy on the artist and composition.

that hypothesis 1 on the perceived aesthetic value of AI-generated art is supported. Similarly, the odds of accuracy increased by 1.27 times for the composition variable, controlling for subjects, 95% CI [1.08, 1.49]. As the coefficient is positive and the log odds are more than 1, this indicates that subjects gave a higher accuracy to human-made paintings. This indicates that hypothesis 2 about the aesthetic appeal of the inclusion of the RoT is also supported by the findings.

However, as previously mentioned in the methods section, the dataset used for the aforementioned tests includes subjects who either had an idea that the data included AI paintings, or they knew one or more human painters, whose paintings are included in the dataset. The exact same tests as above were run, to make sure that these subjects did not cause bias in the data. The output for the artist variable ($B = 0.64$, $SE = .091$, $p < .001$) and composition variable ($B = 0.21$, $SE = .091$, $p = .02$) still resulted in the same overall conclusions, which are that paintings that were human-made and paintings that followed the RoT received a higher accuracy.

5.2 *Response Times*

There are also recorded response times. A Wilcoxon signed-rank test was used to evaluate the differences in response times between paintings made by human artists and an AI. This analysis included 67 pairs of observations. The results from the test indicated a significant difference between the response times for paintings made by a human artist and an AI artist ($W =$

324243, $p < .001$, $r = .15$). Further descriptive statistics demonstrated that human-made paintings caused a median response time of 1.32 seconds ($SD = 0.91$), while AI-generated paintings have a median response time of 1.46 seconds ($SD = 0.90$). This implies that subjects took a significantly longer time to evaluate the human-made paintings.

The same Wilcoxon signed-rank test was used to test between paintings that included the RoT and those that did not include the RoT. This test showed no significant results between the two groups ($W = 383952$, $p = .52$, $r = .018$).

6 DISCUSSION

This study aims to provide further insight into the question of whether AI-generated art is evaluated to be as aesthetically pleasing as human-made art. This thesis narrows down the focus on the issue of the evaluation of aesthetics without the bias that is present when subjects become aware that they are looking at artificially generated art. The experiment that was conducted therefore did not include any suggestion of AI-generated art. This meant that the subjects evaluated the paintings solely on their subjective assessment of the aesthetics. This study also looks into the issue of the aesthetic appeal of the Rule-of-Thirds (RoT). Although this is a popular technique in creating a pleasing spatial composition, there is debate in the literature on whether this effect actually impacts the aesthetic evaluation.

6.1 *Findings*

6.1.1 *Accuracy*

In this section, the findings will be presented and interpreted. The first finding revolved around the accuracy of the independent variable artist. The goal of this test was to determine whether there was a significant difference in the aesthetic evaluation of paintings made by humans versus paintings made by AI. The findings demonstrate that paintings made by AI were significantly rated as less aesthetically appealing in comparison to the paintings made by human artists. With this finding, the hypothesis that human-made art would be evaluated as more aesthetic is supported.

These findings are consistent with other work in the field, which has shown a difference between artwork produced by humans and AI in terms of perceived aesthetic value (Chamberlain et al., 2018). The findings of the present study contribute to the body of literature that contends that people still prefer artwork created by human artists over that created by

AI algorithms (Chamberlain et al., 2018). As a result, since AI-generated art lacks in the area of visual appeal, it could not arouse the same level of aesthetic appreciation.

The findings also shed new light on the prejudice towards AI artwork in the art world that was found in recent literature (Ragot et al., 2020). This prejudice may be caused by worries about the underappreciation of the time and effort that humans put into art, apprehensions about losing one's employment, or the belief that AI simply lacks the authenticity and emotional and intellectual depth attributed to human creativity.

There is still widespread doubt and opposition to adopting AI-generated artwork as a valid art form, despite the major improvements in AI-generated art and the rising acceptance of AI as a creative tool. This study provides new insights that it might not just be the prejudice against AI artwork that is affecting the lower aesthetic appreciation. Although recent AI techniques have come far, it seems that there is still room for improvement in the area of aesthetic appeal.

The results of this study add to the ongoing discussion of how AI and human creativity interact. They enforce the need for a nuanced view of AI artworks, that takes into account its qualities and limitations.

6.1.2 *Rule-of-Thirds*

In addition to the main research question, there is also the sub-question related to the aesthetic value of the RoT. The results of this study provide evidence that paintings that follow the RoT are evaluated to be more visually attractive than paintings that do not follow the RoT. These findings contradict the finding of the study by Amirshahi et al. (2014) that proposes that the RoT might not have the effect on the aesthetic evaluation as presumed. However, The results of this study do underline the importance of the rule of thirds in aesthetic evaluations and emphasise its applicability in the field of paintings. This is consistent with art experts who have an outspoken opinion on this matter (Gooch et al., 2001; Mai et al., 2011) and the findings from the study by Koliska and Oh (2023).

It is vital to remember that while the RoT has a considerable impact on aesthetic judgments, it should not be viewed as an unbreakable rule or a restriction on an artist's ability to express themselves. Experimentation, pushing limits, and breaking from established conventions are frequently components of new innovations in the area of art.

Future studies could go more deeply into the psychological processes that underlie the RoT Additionally, examining how the RoT interacts with other compositional principles may shed light on the interactions of aesthetic perception.

6.1.3 *Response Times*

The response times for the subjects were found to be significantly higher on the paintings made by human artists than on the paintings made by AI. The substantially longer response times for paintings made by human artists could imply that participants needed to use more cognitive processing to assess these pieces of art. Human-made paintings frequently have complex compositional elements and creative methods, that take more time and mental work to appreciate and evaluate. On the other side, AI-generated artwork may lack some characteristics or display more recognizable patterns, resulting in quicker judgments.

Moreover, given the long history of human artistic production, the participants' evaluation processes may have been impacted by pre-existing conceptions or expectations about art. Being a relatively new form of art creation, AI-generated paintings could not have aroused as much reflection or cognitive involvement.

6.2 *Limitations*

There may be methodological elements that had an influence on the results. One of these elements could be the choice of art style. For the reason of homogeneity, all paintings that were selected followed the art style of impressionism. Impressionism is an art style that uses broken brushwork and vibrant colour palettes, in order to capture the effects of light. These artists focus more on capturing a short moment in time, as opposed to pursuing a detailed representation (Samu, 2004). Dall·2 was tasked to recreate this art style given a specific subject matter. From the resulting AI paintings, it is visible that a portion of these artworks includes brushstrokes that are comparably bigger than those of Theodore Robinson and Armand Guillaumin. This causes a blurry effect to take over some of these paintings, causing them to be much less detailed. From the results, this could have influenced the relatively negative evaluation of the AI-generated paintings. However, the paintings that showed this effect are still valid representatives of artwork generated by AI, because they received the instructions to recreate paintings from impressionism. A suggestion for a future study could be to explore different art styles in their experiment.

Furthermore, art revolves around emotion and messaging abstract ideas. This experiment narrowed down the subject matter, leaving little to no room for emotional interpretation or the inclusion of abstract values. In addition, subjects were given only a short window of time to respond. This further caused the subject to provide only a surface-level evaluation of the aesthetics, which meant that the subject had a narrow view in assessing the aesthetics. If participants were given a longer time or unlimited time to respond, they might have provided a more critical assessment of the aesthetics. Future studies could develop an experiment that would allow the subject to get a more critical evaluation. Despite this fact, the overwhelming majority of response times were relatively low compared to the time limit, meaning that, overall, participants may not have needed much more time to respond, even without a time limit.

Another point of attention is the characteristics of the paintings. The subject matter was narrowed down to the inclusion of a partially green landscape with nature and tree(s). Despite this, several paintings included elements on top of these, such as bodies of water. Especially in such a homogeneous dataset, this could have influenced the results. Another influential characteristic could be colour saturation or vibrancy. This is not consistent among all the paintings. A follow-up study could include a wider variety of subject matters to include in the paintings.

6.3 *Societal Impact*

According to this study, AI-generated artwork lacks aesthetic quality, compared to human-made art. However, the quality of the creation of art by AI will likely improve, as the techniques develop further. The careers of artists who rely on their originality and talent to produce original works might be affected if the use of AI-generated artwork becomes more popular.

Moreover, concerns surrounding proper crediting and intellectual property are brought up by AI-generated art (Epstein et al., 2020). Questions concerning correct attribution, copyright violations, and the possibility for AI-generated art to be wrongly ascribed to human artists arise if AI systems are able to create artwork that is aesthetically comparable to human-created art. This can cause legal, cultural, and financial impacts on the art world.

Another potential drawback of AI-generated art must be acknowledged, namely the algorithmic biases ingrained in the training data that may unintentionally reinforce current societal biases or sustain imbalances in representation and diversity within the art domain (Norori et al., 2021; Peng et al., 2022). Future studies in the area of AI-generated art could focus on finding solutions to address these biases and the further implications these biases could cause.

However, there are also those who argue that AI is simply a new expression of art (Chatterjee, 2022) They argue that instead of seeing AI as a threat, to see it as a tool for collaboration to strengthen each other's creativity.

In regards to the RoT, for designers, educators, and artists, these results further support that is useful to comprehend the rule of thirds and how it affects aesthetic judgments. Artists may improve the visual appeal and engagement of their paintings by using this compositional approach to their work. By giving students a fundamental concept of visual composition and empowering them to produce more aesthetically attractive work, educators may stress the value of the rule-of-thirds in art instruction.

7 CONCLUSION

In conclusion, this study examined the aesthetic perception of human-made art and AI-generated art, as well as the effect of the Rule-of-Thirds (RoT). The results showed a significant difference in the evaluation between human-made art and AI-generated art. Although recent AI advances have produced considerable results in the creation of art, the findings in this study suggest that art created by humans is more aesthetically appealing to humans. This finding contributes to our understanding of visual aesthetics. In addition, it was found that the RoT plays a significant role in the evaluation of aesthetics in artwork. The findings show that the visual harmony and aesthetic appeal of artworks are improved by the use of the compositional rule, the RoT.

REFERENCES

Amirshahi, S. H., Hayn-Leichsenring, G. U., Denzler, J., & Redies, C. (2014). Evaluating the rule of thirds in photographs and paintings. *Art & perception* **2(1–2)**. https://doi.org/10.1163/22134913-00002024

Chamberlain, R., Mullin, C., Scheerlinck, B., & Wagemans, J. (2018). Putting the art in artificial: Aesthetic responses to computer-generated art. *Psychology of Aesthetics, Creativity, and the Arts* **12(2)**. https://doi.org/10.1037/aca0000136.

Chatterjee, A. (2022). Art in an age of artificial intelligence. *Frontiers in Psychology* **13**. https://doi.org/10.3389/fpsyg.2022.1024449.

CineD. (2023). *Creating entire movies from text using google's ai tools*. Retrieved May 17, 2023, from https://www.cined.com/creating-entire-movies-from-text-using-googles-ai-tools/

Elgammal, A. (2019). AI Is Blurring the Definition of Artist. *American Scientist* **107(1)**. https://doi.org/10.1511/2019.107.1.18

Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1706.07068

Epstein, Z., Levine, S., Rand, D. G., & Rahwan, I. (2020). Who gets credit for ai-generated art? *iScience* **23(9)**. https://doi.org/https://doi.org/10.1016/j.isci.2020.101515.

Eshraghian, J. (2020). Human ownership of artificial creativity. *Nature Machine Intelligence* **2(4)**. https://doi.org/10.1038/s42256-020-0161-x.

Feagin, S. L., & Maynard, P. (1997). *Aesthetics*. New York: Oxford University Press.

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. *Computer Vision and Pattern Recognition*. https://doi.org/10.1109/cvpr.2016.265

Glynn, B. P. (2023). *Sony world photography award 2023: Winner refuses award after revealing ai creation*. https://www.bbc.co.uk/news/entertainment-arts-65296763

Gooch, B., Reinhard, E., Moulding, C., & Shirley, P. (2001). Artistic Composition for Image Creation. In S. J. Gortle & K. Myszkowski (Eds.), *Eurographics workshop on rendering*. The Eurographics Association.

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

Guillaumin, A. (1911). *Gust of wind, le brusc*. Retrieved May 12, 2023, from https://www.armandguillaumin.org/Gust-Of-Wind-Le-Brusc.html

Hannah Kaube, A. E., & Rahman, R. A. (2023). *Separating art from the artist: The effect of negative affective knowledge on erps and aesthetic experience*. https://doi.org/10.1371/journal.pone.0281082

Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2021). *Coin: Conditional inference procedures in a permutation test framework* [R package version 1.4-2]. https://CRAN.R-project.org/package=coin

Koliska, M., & Oh, K. (.-K. (2023). Guided by the grid: Raising attention with the rule of thirds. *Journalism Practice* ***17*****(2)**. https://doi.org/10.1080/17512786.2021.1916402.

Kruger, J., Wirtz, D., Van Boven, L., & Altermatt, T. (2004). The effort heuristic. *Journal of Experimental Social Psychology* ***40*****(1)**. https://doi.org/https://doi.org/10.1016/S0022-1031(03)00065-9.

Mai, L., Le, H. M., Niu, Y., & Liu, F. (2011). Rule of thirds detection from photograph. *International Symposium on Multimedia*. https://doi.org/10.1109/ism.2011.23.

Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., & Tzovara, A. (2021). Addressing bias in big data and ai for health care: A call for open science. *Patterns* **2(10)**. https://doi.org/10.1016/j.patter.2021.100347

Obrador, P., Schmidt-Hackenberg, L., & Oliver, N. (2010). The role of image composition in image aesthetics. *2010 IEEE International Conference on Image Processing*. https://doi.org/10.1109/ICIP.2010.5654231.

OpenAI. (2022). *DALL-E 2*. Retrieved May 15, 2022, from https://openai.com/dall-e-2/

Peng, A., Nushi, B., Kiciman, E., Inkpen, K., & Kamar, E. (2022). Investigations of performance and bias in human-ai teamwork in hiring. *Proceedings of the AAAI Conference on Artificial Intelligence* ***36*****(11)**. https://doi.org/10.1609/aaai.v36i11.21468

Portrait by ai program sells for $432,000 [Retrieved may 14, 2023, from]. (2018). https://www.bbc.com/news/technology-45980863

R Core Team. (2020). *R: A language and environment for statistical computing* [version 4.0.3]. R Foundation for Statistical Computing. https://www.R-project.org/

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J. A., Krueger, G., & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *arXiv (Cornell University)*. http://export.arxiv.org/pdf/2103.00020

Ragot, M., Martin, N., & Cojean, S. (2020). Ai-generated vs. human artworks. a perception bias towards artificial intelligence? *Extended Abstracts* *of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3334480.3382892.

Robinson, T. (1890). *Olive grove, capri*. Retrieved May 12, 2023, from https://www.wikiart.org/en/theodore-robinson/olive-grove-capri-1890

Samu, M. (2004). *Impressionism: Art and modernity*. The Metropolitan Museum of Art.

Tan, X., & Li, X. (2021). A tutorial on ai music composition. *Proceedings of the 29th ACM International Conference on Multimedia*. https://doi.org/10.1145/3474085.3478875.

Tang, A., Jiang, W., Zhang, J., Wang, S., & Gao, Y. (2020). Composition harmony assessment method based on rule of thirds. *2020 International Conference on Culture-oriented Science & Technology (ICCST)*. https://doi.org/10.1109/iccst50977.2020.00034.

Wickham, H. (2020). *Tidyr: Tidy messy data* [R package version 1.1.2]. https://CRAN.R-project.org/package=tidyr

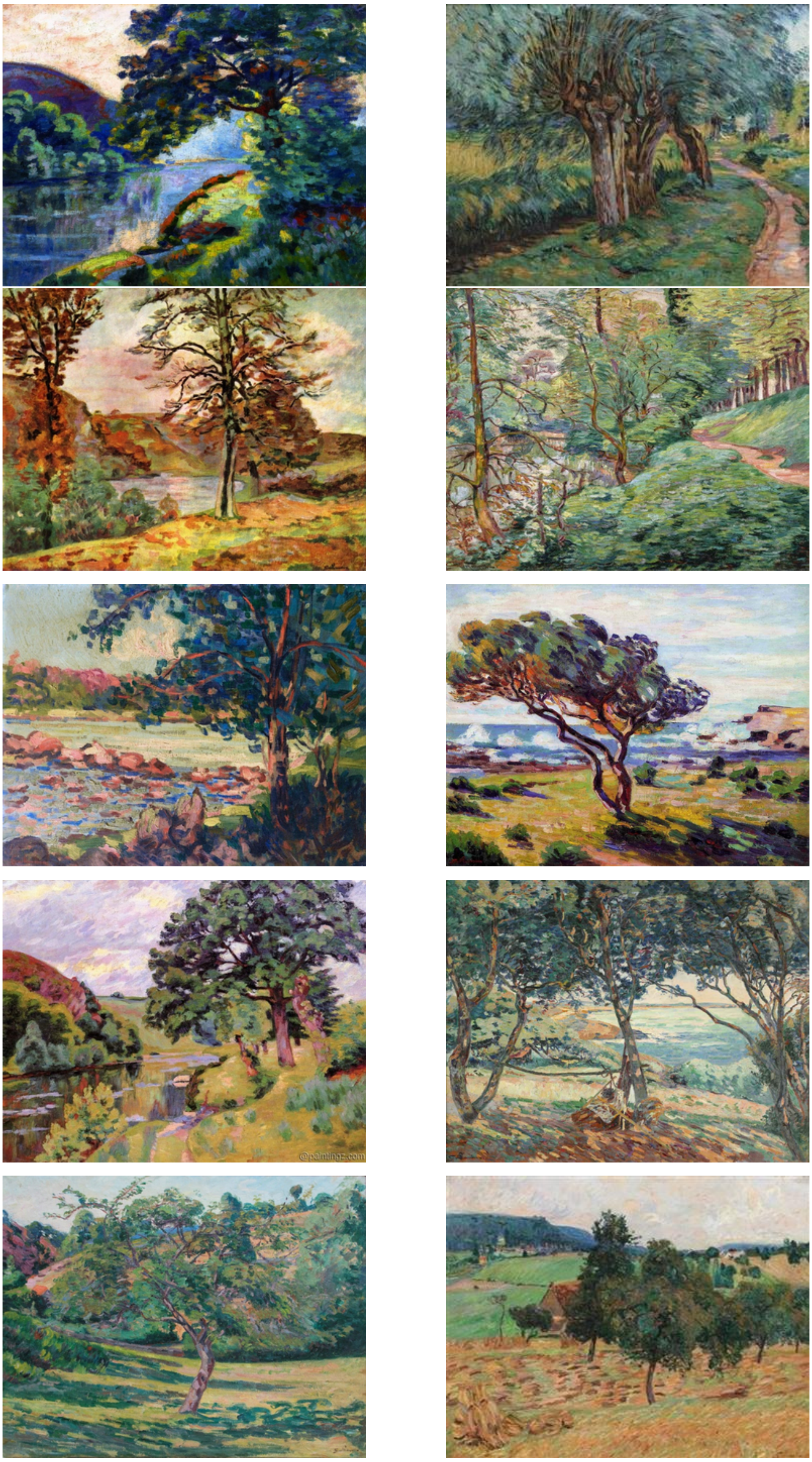
Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation* [R package version 1.0.4]. https://CRAN.R-project.org/package=dplyr

Zaidel, D. W. (2010). Art and brain: Insights from neuropsychology, biology and evolution. journal of anatomy. *Journal of anatomy* ***216*****(2)**. https://doi.org/10.1111/j.1469-7580.2009.01099.x

Appendix

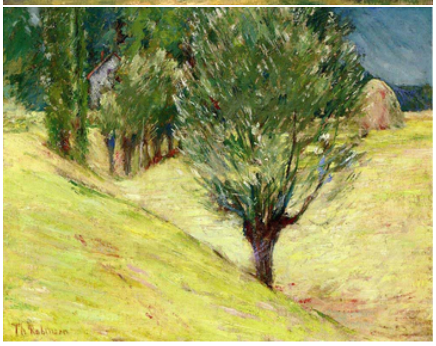
These paintings are made by the artist Armand Guillaumin.

RoT | no RoT



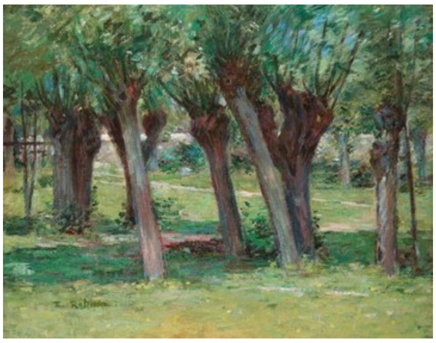
These paintings are made by the artist Theodore Robinson.

RoT


no RoT





These paintings are made by Dall·E 2 (OpenAI, 2022), using the prompt: "An impressionist oil painting of a landscape with trees"

RoT　　　　　　no RoT

These paintings are made by Dall·E 2 (OpenAI, 2022), using the prompt: "An impressionist oil painting of a landscape with trees"

RoT | no RoT