The Influence of Creator Information on Preference for Artificial Intelligence– and Human–generated Artworks

Seungmin Nam¹ · Jiwon Song² · Chai-Youn Kim³†

Abstract

Purpose: Researchers have shown that aesthetic judgments of artworks depend on contexts, such as the authenticity of an artwork (Newman & Bloom, 2011) and an artwork’s location of display (Kirk et al., 2009; Silveira et al., 2015). The present study aims to examine whether contextual information related to the creator, such as whether an artwork was created by a human or artificial intelligence (AI), influences viewers’ preference judgments of an artwork. Methods: Images of Impressionist landscape paintings were selected as human-made artworks. AI-made artwork stimuli were created using Google’s Deep Dream Generator by mimicking the Impressionist style via deep learning algorithms. Participants performed a preference rating task on each of the 108 artwork stimuli accompanied by one of the two creator labels. After this task, an art experience questionnaire (AEQ) was given to participants to examine whether individual differences in art experience influence their preference judgments. Results: Setting AEQ scores as a covariate in a two-way ANCOVA analysis, the stimuli with the human-made context were preferred over the stimuli with the AI-made context. Regarding the types of stimuli, the viewers preferred AI-made stimuli to human-made stimuli. There was no interaction effect between the two factors. Conclusion: These results suggest that preferences for visual artworks are influenced by the contextual information of the creator when the individual differences in art experience are controlled.

Key words: Art, Ai, Creator, Context, Preference, Art Experience, Individual Difference

1. INTRODUCTION

Appreciation of an artwork requires a sequence of processes ranging from sensory analyses to cognitive and affective processes with socio-cultural influences (Leder et al., 2004). The model of aesthetic experience Leder has proposed (and later updated) provides a framework for aesthetic experiences including perception of visual artworks to aesthetic judgment (Leder et al., 2004; Leder & Nadal, 2014). The model begins with analyses of a visual artwork in terms of visual features such as contrast, complexity, or symmetry. These analyzed features of the artwork are put together with pre-existing memory or personal experience for explicit classification of its content and style. The cognitive mastering stage is followed in which the artwork is interpreted as artistic- or self-related meaning. With the interpretation, aesthetic judgment and aesthetic emotion

† This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A5A2A01023762).

¹ Seungmin Nam: Master, School of Psychology, Korea University
² Jiwon Song: Ph.D. candidate, School of Psychology, Korea University
³† (Corresponding Author) Chai-Youn Kim: Professor, School of Psychology, Korea University / E-mail: chaikim@korea.ac.kr / TEL: 02-3290-2866
are finally derived from the cognitive and affective states.

Another factor Leder’s model pays attention to is the context in which an artwork is appreciated. There have been a few studies demonstrating that aesthetic judgments vary depending on the contextual information such as the authenticity (Huang et al., 2011; Newman & Bloom, 2012), the source (Kirk et al., 2009; Silveira et al., 2015), or the title of an artwork (Leder et al., 2006). Another critical contextual information is the information of the artist who created the artwork. Most of us have an experience of viewing an artwork drawn with unsophisticated primary colors with basic shapes on which our first impression was not very special, yet knowing that it is a work of Paul Klee makes the work look great. This is an instance where the contextual information of the creator comes into play in our aesthetic judgment of the work.

A series of studies have shown that an artwork labelled with ‘created by a famous artist’ is evaluated as more creative and valuable compared to ‘created by an amateur artist’ (Bernberg, 1953; Gergen & Breger, 1965). This effect has been called as prestige bias effect. In contrast, an opposite effect, dubbed ‘the reversal prestige bias effect’, was observed in a more recent study. Isham and colleagues explored this effect that the artworks assigned to youth artists were evaluated higher compared to the ones assigned to famous artists (Isham et al., 2010). For the stimuli in the ‘youth artists’ context, the viewers would have more generous criteria of aesthetic evaluation.

The choice of the present study concerning the creator, however, is neither a particular individual such as Paul Klee nor depending on the fame. We took AI into consideration as a new agent of art creation in the present study. Our choice of the AI creator, in comparison with a human creator, is not only timely and interesting but also has a couple of methodological advantages in an investigation into the contextual effect of the creator information on aesthetic judgments. First, the information of the ‘AI creator’ is not limited to viewers’ prior understanding of the creator. If a particular human artist’s name is given as contextual information accompanying an artwork, how much the viewer knows about or prefers the artist will affect heavily his/her aesthetic judgment. Rather than providing one or two specific human artists’ names as the creator information, this study provides ‘AI-made’ information in contrast to ‘human-made’ information. Therefore, the contextual information regarding the creator (i.e., ‘human-made’) was not based on the prestige bias. Nor is it based on the reversal prestige bias since the AI is not necessarily inferior in making art compared to any human being other than professional artists. AI can be rather viewed as a novel type of art creator irrespective of relative superiority/inferiority to human. Another methodological advantage of considering the AI creator is that the physical differences between the AI-made artworks and human-made artworks can be minimized (Elgammal et al., 2017), which would be expected to maximize the room for the aesthetic judgment moderated by the contextual information. By utilizing an AI program mimicking the style of a given artwork created by a human artist, the similarity in terms of the style and quality between the AI- and human-made artworks has been secured relatively.

Up to now, there have been only a limited number of studies investigating whether information on the attribution of the AI creators can bias viewers’ aesthetic evaluations, which demonstrated seemingly inconsistent results. In the study of Hong & Curran (2019), information on the attribution of the creators (i.e., AI or human) did not have a significant effect on the appraiser’s evaluation, whereas in the study of Ragot and colleagues (2020), the ‘AI’ creator information induced lower aesthetic evaluation compared to the ‘human’ information. Furthermore, Gangadharbatla revealed that the attribution knowledge – ‘AI’ or ‘human-made’ – is related to the type of a visual artwork (Gangadharbatla, 2021).
The Influence of Creator Information on Preference for Artificial Intelligence- and Human-generated Artworks

Such inconsistencies of the previous results might not only be related to the characteristics of the stimuli but also the characteristics of the beholder. According to the model of Leder’s, an individual’s previous experience and expertise in art are factors that influence the overall art appreciation process (Leder et al., 2004; Leder & Nadal, 2014). Therefore, in the present study, we sought to investigate whether viewers’ aesthetic preference of an artwork with contextual information of the creator is related to their prior aesthetic experiences. In addition, we hypothesized that the impression of the creator context – ‘AI’ or ‘human’ – might be shaped depending on viewers’ personality traits (Yoon & Lee, 2016). According to the study by Feist & Brady (2004), the openness personality trait of the viewer correlated positively with their preference of abstract over representational art. Compared to traditional representational art, abstract art is a relatively recent and more unfamiliar artistic style. Open participants who can accept unfamiliar and novel concepts showed a high preference for novel abstract art (Feist & Brady, 2004). In a similar vein, breaking away from the traditional way of thinking that only humans can create art, the concept of AI creator is emerging in art. Therefore, it is probable that aesthetic evaluation of novel concept is closely related to openness. Based on this previous finding, we aimed to test whether the viewers of higher openness might prefer the artwork labeled as ‘AI-made’ more.

To sum up, the present study investigated whether the contextual information of AI or human creator affects viewer’s art appreciation and preference. By utilizing Google’s Deep Style program implementing deep learning algorithms, artwork stimuli were made by AI mimicking the style of a human artist’s piece of work. Lastly, we examined whether individual differences of art experience and openness as a primary personality trait affect aesthetic preference under the contextual modulation. We predicted that the creator information would influence the viewers’ aesthetic preference, while the actual stimulus difference created by AI and human would be minimal, and therefore, have little impact on aesthetic preference. We also predicted that individual differences in art experience and openness would be another important factor for aesthetic preference.

2. METHOD

2.1. Participants

A total of 70 individuals participated in the study; 30 individuals participated in the preliminary experiment for stimulus selection (19 females/11 males, 20-28 years of age), and 40 individuals participated in the main experiment (18 females/22 males, 20-29 years of age). We conducted power analysis with G*Power 3 software (Faul et al., 2007) using an alpha of .05, a power of .80 and a medium effect size ($\eta_p^2 = .06$). The minimum sample size required to test was 23 participants. We included the even larger number of participants in consideration of individual variances in aesthetic preference judgments. All of them were undergraduate or graduate students at Korea University or two other local universities. All participants had a normal or corrected-to-normal vision. They received no formal training in art, nor did they hold a degree in an art-related major. All of them gave informed consent approved by the Korea University Institutional Review Board (1040548-KU-IRB-16-199-A-1).

2.2. Stimuli

In an attempt to use artwork images as visual stimuli, we selected human-made artwork images first and then generated AI-made artwork images. To control potential differences of preference for representational from abstract arts (Flexas et al., 2014; Mastandrea et al., 2009) all human-made artwork images were limited to 19th-century Impressionist landscape paintings that do not include human figures. Twenty landscape images were selected from
each of the six Impressionist artists including Claude Monet, Pierre-Auguste Renoir, Vincent van Gogh, Paul Cezanne, Camille Pissarro, and Paul Gauguin (http://www.wikiart.org/), resulting in a total of 120 images. The artist’s signature, if included, was removed by using Adobe Photoshop software. To generate AI-made artwork images, we utilized Google’s Deep Dream Generator (https://deepdreamgenerator.com/) including the Deep Dream and the Deep Style programs. Deep Dream Generator is an implementation of Convolutional Neural Network for an artificial system to learn a style of an input image via machine learning algorithms (Mordvintsev et al., 2015; Szegedy et al., 2015). We specifically used the Deep Style program, which is capable of learning a style of a given input and using it to convert a new image following the style. To run this program, therefore, a pair of images – i.e., an input image for style learning and another image for conversion – is necessary. The Deep Style program was used to generate 120 AI-made artwork images by converting 120 landscape photograph images into images with the artistic style learned from each of the 120 human-made artwork images (Fig. 1). As a result, a total of 240 images (120 AI-made and 120 human-made) were used for the preliminary experiment for stimulus selection.

Although AI-made painting images generated by Deep Style were designed to resemble human-made artwork images, there might be some visual characteristics distinguishing the AI-made painting images from the human-made artwork images stemming from how Deep Style works. In order to maximize the potential contextual influence of the creator information on the aesthetic preference in the main experiment by minimizing the feature differences between the two classes of the images, we conducted a preliminary experiment. The experiment was conducted in a dark room where the computer monitor was the only source of illumination. Images were presented on a 19-inch CRT monitor (1024 x 768 resolution, 60-Hz frame rate, 43-cm distance) using Matlab 9.1 (MathWorks, Inc., MA) with Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997). The size of each image was differentially adjusted for its ratio of width and height, and less than or equal to a maximum of 17.5° x 15° in visual angle. They were presented on a light gray background at the center of the monitor. Participants sat in front of the monitor while their head positions were stabilized using a head and chin rest. Thirty participants were asked to rate each of the 240 images based on their degree of confidence regarding its creator using a 7-point Likert scale (1: ‘It was definitely created by a human artist.’ to 7: ‘It was definitely created by an AI creator’, with 4 being ‘not sure’ response). To exclude the confounds of fame and familiarity for paintings, as well as to maximize the potential contextual influence of the creator information, the images with moderate rating scores were selected. For instance, the images rated extremely low are likely to be familiar Impressionist paintings to the participants, whereas the images rated extremely high might have some qualities unlikely to be created by human. Based on the rating results, we selected the images of which mean rating scores ranging between 2.5 and 5.5 with standard deviation smaller than the mean standard deviation of the entire images.

As a result, there were 68 pairs of AI- and hu-
man-made artwork images, and 20 AI-made and 20 human-made artwork images that were not matched with each other. Therefore, a total of 176 artwork images were selected as the stimuli for the main experiment.

2.3. Procedure

Among the selected 176 stimuli, 136 stimuli were the matched pairs (i.e., AI-made stimuli following the style of the corresponding human-made stimuli) and the two matched stimuli of a pair retained high similarity. To control the amount of exposure to similar stimuli across the participants that might influence participants’ aesthetic preference (Song et al., 2021) we decided to present only a half of the pairs – i.e., 68 paired stimuli (34 human-made and 34 AI-made) to one half of the participants and the other 68 paired stimuli to the other half of the participants. All of the 40 non-matched stimuli were presented to the entire participants. Therefore, a total of 108 stimuli (54 human-made and 54 AI-made) were presented to each of the participants. Among the 54 human-made stimuli, a half was presented in the ‘human-made’ context and the other half was presented in the ‘AI-made’ context. Likewise, 54 AI-made stimuli were presented either in the ‘human-made’ or in the ‘AI-made’ context.

A 2 x 2 factorial block design was used where the stimulus and the context were systematically manipulated. Each block was assigned to one of the four experimental conditions: AI-made stimulus with ‘AI-made’ context (AA), human-made stimulus with ‘AI-made’ context (AH), AI-made stimulus with ‘human-made’ context (HA), and human-made stimulus with ‘human-made’ context (HH).

As shown in Fig. 2, each block consisted of three trials. On each trial, the creator information (i.e. ‘This is an AI-made artwork’ or ‘This is a human made artwork’) was presented for 2 seconds, and then, the painting image was presented for 4 seconds. Subsequently, participants were asked to rate how much they preferred the painting image on a 4-point Likert scale (1: ‘I definitely do not prefer it’ to 4: ‘I definitely prefer it’).

After the main experiment, two surveys were conducted to assess individual differences among the participants. One was the Art Experience Questionnaire (AEQ, Chatterjee et al., 2010) that includes the questions of the personal history of art education and the time devoted to the art-related activities. The AEQ was included to examine whether the degree of an individual’s experience in art affects one’s preference for artworks with creator information. The other was the survey items related to openness to experience sorted from the HEXACO-PI (http://hexaco.org, Lee & Ashton, 2004) including four sub-scale measures: aesthetic appreciation, inquisitiveness, creativity, and unconventionality. The survey was to take into consideration the possibility that the preference for AI-made artworks is related to openness to experience among an individual’s personality factors. It took 5 minutes on average to complete the two surveys.

3. RESULTS

3.1. Preference ratings

Fig. 3a shows the mean preference rating scores for the four experimental conditions. The mean preference rating scores were 2.77 (±1 standard error of the mean; ISEM = .06) for AA, 2.42 (ISEM = .06) for AH, 2.81 (ISEM = .06) for HA, and 2.49 (ISEM = .06) for HH conditions, respectively. A two-way repeated measures
analysis of variance (ANOVA) showed no statistically significant main effect of context \((F(1,35) = 1.17, p = .30)\) nor the two-way interaction between context and stimulus \((F(1,35) = .16, p = .70)\). In contrast, the main effect of stimulus was statistically significant in that the mean preference rating score for \textit{AI-made} stimuli was higher than the mean preference rating score for \textit{human-made} stimuli \((F(1,35) = 79.33, p < .001)\).

### 3.2. Correlations between preference ratings and the post-task surveys

To examine whether individual differences had an impact on the preference for the artworks, we conducted correlation analyses for the mean preference rating scores and post-task questionnaires including the AEQ and the openness scores from the HEXACO-PI. Fig. 4a shows that individual AEQ scores negatively correlated with mean preference scores for the stimuli shown in the \textit{‘human-made’} context (Person’s \(r = -.36, p = .03\)). There was no statistically significant correlation between AEQ scores and mean preference scores for the stimuli shown in the \textit{‘AI-made’} context. Among the three subscales of the AEQ, only duration of art activity negatively correlated with the mean preference scores for the stimuli shown in the \textit{‘AI-made’} context. Among the three subscales of the AEQ, only duration of art activity negatively correlated with the mean preference scores for the stimuli shown in the \textit{‘human-made’} context (Person’s \(r = -.46, p = .005\), shown in Fig. 4b). In contrast, individual openness scores from the HEXACO-PI survey did not show any statistically significant correlations with the preference rating scores \((r = -.29, p = .09\) for the preference rating scores for stimuli in the \textit{‘AI-made’} context; \(r = -.23, p = .18\) for the preference rating scores for stimuli in the \textit{‘human-made’} context).

### 3.3. Preference ratings with individual differences factored out

Accordingly, we considered the AEQ score as a factor affecting the preference of an artwork under a certain creator context. To examine the effect of contextual information of the creator on aesthetic preference of an artwork, we factored the AEQ score out as a covariate by conducting a two-way analysis of covariance (ANCOVA) with context and stimulus as the two factors. When taking the overall AEQ score out as a covariate, the preference rating score for the \textit{‘human-made’} context tended to be significantly higher than in the \textit{‘AI-made’} context \((F(1,35) = 3.97, p = .054)\). In the stimulus aspect, \textit{AI-made} stimuli were significantly more preferred to \textit{human-made} stimuli \((F(1,35) = 28.46, p < .001)\). In the aspects of context and stimulus, however, there was no significant two-way interaction effect \((F(1,35) = .08, p = .78)\). When taking durations of art activity out as a covariate, both the main effects of context and stimulus were statistically significant \((F(1,35) = 8.65, p < .01)\) for the main effect of context; \((F(1,35) = 44.92, p < .001)\) for the main effect of stimulus, respectively). There was no interaction effect between the two factors \((F(1, 35) = .18, p = .67)\).
4. DISCUSSION

Using both AI- and human-made stimuli, the current study investigated a factor important for aesthetic preference judgment of an artwork - i.e., the contextual information on the creator. Importantly, the current results showed that individual differences in the degree of art experience derived by the AEQ moderated preference of a visual artwork in the ‘creator’ context. The more experiences a viewer has in art, the less he/she prefers an artwork in the ‘human-made’ context. To examine the genuine contextual effect without the intervention of individual differences in art, we took individual’s overall AEQ scores out while analyzing the preference rating scores under the experimental conditions. The results showed that artworks were preferred more in the ‘human-made’ context than in the ‘AI-made’ context. For the stimulus factor, however, human-made stimuli were less preferred to AI-made stimuli.

These results are in line with some of our predictions, but not all. Unlike our prediction, the openness personality trait did not show any effect on the contextual effect of the creator information nor aesthetic preference of the artworks. However, the current results provided supporting evidence for the interplay between the contextual effect of creator information on aesthetic preference of an artwork and the viewer’s art experience. This novel finding might explain the inconsistencies found in the limited number of studies regarding the effect of the ‘AI creator’ information on viewers’ aesthetic evaluations. As reviewed in Introduction, Hong & Curran (2019) showed no effect of the information of the creators (i.e., ‘AI’ or ‘human’) on the viewers’ evaluation of the artwork. Our results are in line with the results from Hong and Curran when the individual differences were not considered. In contrast, our results become consistent with the results from Ragot and colleagues (2020) when the individual differences in art experience were considered; The creator information affected the viewers’ aesthetic preference of an artwork. Furthermore, viewers gave higher preference ratings for the artworks with the ‘human-made’ label than for the artworks with the ‘AI-made’ label. Taken together, our results highlight the importance of considering the individual differences factor of the beholder in the studies of aesthetic appreciation.

Especially, the results of the correlation analyses between the AEQ scores and the preference ratings in the human-made context made us consider the individual differences in art experience are important factor in the contextual effect on art appreciation. Before factoring out the AEQ scores, the main effect of the context was not shown, as the preference ratings seemed not to be affected by the contextual information - AI made or Human-made. After controlling the AEQ scores as the covariate variable, however, the preference ratings on the context of ‘human-made’ artwork tended to be significantly higher than the context of ‘AI-made’ artwork. These results suggest that the creator information influences the viewer's art appreciation, which is highly involved by its previous art experience. It was also noted that individual differences of viewers, such as previous art experiences, is a convincing factor that should be considered in the studies of aesthetic appreciation. A further study is needed to scrutinize whether individual characteristics mediate the preference of an artwork along with its creator information.

Perhaps the most surprising results from the current study contrary to our prediction was the difference in preference based on the stimulus difference between AI- and human-made artworks. Despite the physical similarity between the AI-made and human-made stimuli, AI-made stimuli were rated higher in preference judgment task than human-made stimuli. There is no reason to take such preference for AI-made artworks as evidence of the participants’ awareness of the actual creator or the intrinsic preference for AI-made artworks. Rather, it might be related to the way the AI-made artwork images were created. As explained in the Methods section in detail, a landscape photograph was used as an input.
to the Deep Style program, which was converted mimicking the style of a particular human-made landscape painting. Therefore, the layout of the scene and the composition of the AI-made artwork stimuli retain the typicality and familiarity of the ordinary scene layout of the photographs, which might be a reason behind preference for those stimuli as both typicality (Boselie, 1991; Hekkert & van Wieringen, 1990; Martindale & Moore, 1988) and familiarity (Berlyne, 1970; Dearden, 1984; Song et al., 2021) affect aesthetic judgments.

This last point bears significance in terms of the participants in the current study. We recruited individuals with no formal training in art, nor a degree in an art-related major. The preference towards more typical, familiar landscape composition in AI-made artworks might be related to the characteristics of those participants. Studies have suggested that experts value perceptual challenge, ambiguity, and novelty more in appreciation of art (Belke et al., 2015; Jakesch & Leder, 2009; Muth & Carbon, 2013). Therefore, it will be worth testing art experts in a future study to examine whether experts’ preference also differ between AI- and human-made artworks in the same direction of the current, non-experts’ results. Testing art experts will also add to the current findings demonstrating that individuals with greater art experience tend to judge the artworks with the ‘human-made’ label more negatively compared to those with less art experience, although none of them was experts in art.

Last but not least, a future study should consider different types of AI-made artwork stimuli. In the current study, we intentionally chose Deep Style program to mimic particular kinds of styles of the human artists using landscape photographs as inputs to the AI. We also limited our human-made stimuli to the 19th-century Impressionist landscape paintings. Though this was on purpose to control a potential difference of preference for representational from abstract arts (Flexas et al., 2014; Mastandrea et al., 2009), it might have been one of the driving factors of the current findings. The contextual information of the AI creator has been shown to leave a greater impact on appreciation of abstract than representational arts as well (Gangadharbatla, 2021). As greater individual variance of preference and contextual modulation are expected, employing more abstract, unusual, and original artistic creations of AI might generalize the current findings. This will also extend our understanding and discussion over AI as an agency of art creation and how viewers embrace AI’s creative outcomes.

REFERENCES


Silveira, S., Fehse, K., Vedder, A., Elvers, K., & Hennig-Fastm K. (2015). Is it the picture or is it
the frame? An fMRI study on the neurobiology of framing effects. *Frontiers in Human Neuroscience*, 9, 528.


Received: 2022.05.26

Revised: 2022.06.27

Accepted: 2022.06.30