



Are you ready for artificial Mozart and Skrillex? An experiment testing expectancy violation theory and AI music

new media & society
1–16

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1461444820925798

journals.sagepub.com/home/nms



Joo-Wha Hong 

Qiyao Peng

Dmitri Williams

USC Annenberg School for Communication and Journalism, USA

Abstract

This study employs an experiment to test assessments of music composed by artificial intelligence. We examined the influence of (a) met or unmet expectations about artificial intelligence (AI)-composed music, (b) whether the music is better or worse than expected, and (c) the genre of the evaluation of music using a 2 (expectancy violation vs confirmation) × 2 (positive vs negative evaluation) × 2 (electronic dance music vs classical) design. The relationship between the belief about creative AI and the music evaluation was also analyzed. Participants ($n = 299$) in an online survey listened to a randomly assigned music piece. The acceptance of creative AI was found to have a positive relationship with the assessment of AI-composed music. A two-way interaction between the expectancy violation and its valence, and a three-way interaction between the expectancy violation, its valence, and the genre of music were found. Implications for Expectancy Violation Theory and AI applications are discussed.

Keywords

Artificial intelligence, creative AI, creativity, expectation violation theory, human–computer interaction, music

Corresponding author:

Joo-Wha Hong, USC Annenberg School for Communication and Journalism, 3502 Watt Way, Los Angeles, CA 90089, USA.

Email: joowhaho@usc.edu

Artificial intelligence (AI) has permeated our lives in various ways, such as self-driving cars and personal AI assistants (Adams, 2017). As people are exhibiting an increasing interest in AI, they have expectations and concerns with how these new technologies will impact society (Shermer, 2017). Some people are concerned that AI will outperform humans, which is reasonable given recent cases of computers surpassing humans, such as an AI Go player defeating human champions (Curran et al., 2019). However, while it is becoming expected for AI to outperform humans in tasks that require mathematical calculations, it is still questionable whether AI can do the same in creative and artistic pursuits. This is not a thought exercise, as there have been long-standing efforts to understand human creativity in order to develop creative AI based on the presumption that AI can imitate human processes (Boden, 1998).

To determine whether AI can be creative, its own creations inevitably will be judged on its artistic merits. Coeckelbergh (2017) addressed this issue by saying that its creations should be considered artwork based on both subjective and objective criteria. He argued that if there are objective criteria to be fulfilled to be called art, then AI can easily produce products that meet those criteria. If the definition of art is subjective, then everything, including AI-created products, has an equal chance to become artwork. However, even though it is possible to define AI-created works as art, whether people appreciate them as such is a different question. We do not have a clear answer yet, but there are already AI programs that create their own works, including music. Recently, a music-composing AI has been granted copyrights of its own soundtracks (Hewahi et al., 2019). Commercially available, and perhaps popular, AI music may arrive soon.

AI and music

The effort to create AI that performs musical tasks began even before the term “artificial intelligence” was familiar to the public (Meehan, 1979; Roads, 1980, 1985), and the field has grown enough to be called Artificial Intelligence and Music (AIM) (Camurri, 1990). AI is expected to influence music in terms of not only composing, but also performance, and even music education (Zulić, 2019). However, most research has focused on AI’s musical cognition or copyright issues (Bharucha and Olney, 1989; Longuet-Higgins, 1994; Stere and Truşan-Matu, 2017; Sturm et al., 2019), and AI that can listen to music and interact with performers simultaneously (Baird et al., 1993). Zulić (2019) argues that it is now clear that the intrusion of AI in music composing is inevitable, so artists should decide between thinking of AI as a threatening enemy or a future collaborator. While AI-created music is getting closer to its public moment, it is not clear if the human audience is ready. Their attitudes will influence the AI music industry and new musical trends in production and consumptions may depend on people’s reactions.

AI-composers create music by analyzing 30,000 music scores and then creating their own music pieces based on an interpretation using a mathematical model (Barreau, 2018). Some people may say that this is not purely creative since it is more of an imitation of previous music pieces (Jennings, 2010). However, it should be acknowledged that people also learn to make something unique by imitating previous works (Jackson, 2017; Turkle, 2005). If the procedure of creative thinking is the same, then rejecting AI’s creativity comes only from the thought that an AI is not human. The question is whether

people are willing to accept this, yet there has been no empirical research into human reactions and perceptions of AI-created music.

Perceptions of creative AI

It is expected that study of creative AI is a contribution to the field of human–machine communication (HMC) in terms of metaphysical implication, which refers to obscuring boundaries between humans and machines, as creativity is often regarded as an attribute that only belongs to humans (Guzman and Lewis, 2020). Because music is not the only medium where AI creations and reactions can be studied, we can look to other creative fields for parallels. The most studied has been with AI-created visual art. Google’s DeepDream creates hallucinogenic images with its neural network (Marzano and Novembre, 2017). Creative Adversarial Networks (CAN) is another image-creating AI program that creates art by “maximizing deviation from established styles and minimizing deviation from art distribution” (Elgammal et al., 2017: 2). AI is also challenging the movie industry. “Sunspring” is a short movie based on a script by an AI program using neural networks, and “Zone Out” is a movie directed by an AI program (Furness, 2016; Goode, 2018). Multiple experimental studies have been conducted to test how people perceive AI-created artwork. One central finding has been that people’s perception of AI art is positively related to their belief about the ability of AI to be creative (Chamberlain et al., 2018; Hong and Curran, 2019). In other words, a predisposition to be open to AI predicts enjoyment of its products. These results indicate that the evaluation of artwork is biased by perceptions of AI rather than the quality of the art itself. Therefore, it follows that people with the belief that AI can be creative will appreciate music composed by AI more. In turn, this belief appears to be based on a more general set of knowledge and attitudes about AI in general. Sundar and Kim (2019) found that people expect machines to be more just and trustworthy than humans through what they dub “machine heuristics.” In psychology, heuristics refers to efficient cognitive processes that help people to make quick judgments with less effort using what they learned from cultural, institutional, and ideological experiences (Gigerenzer and Brighton, 2009; Gigerenzer and Gaissmaier, 2011; Petersen, 2012). Therefore, machine heuristics are mental shortcuts people use based on their preexisting perceptions of machines when processing machine-related information. It is found from previous studies that people expect AI to have distinctive attributes from humans. For instance, a study found that people saw news articles written by an algorithm as more objective and inducing less emotional involvement than human-written ones because of machine heuristics (Liu and Wei, 2019). Another study found that people consider AI to be less autonomous than humans (Hong and Williams, 2019). This perception may influence the evaluation of AI-created music, since people may have different thoughts on whether the autonomy of the creator is a requirement for producing art. If people think creativity is an innate characteristic that is only possible or acceptable in humans, then AI-composed music would be devalued because of its source. Even though a machine’s performance is indistinguishable from that of humans, there are still people who think that it is not “humanlike” (McCarthy, 2007). Because people distinguish AI from humans by having various machine

heuristics, how people evaluate composed music will also be different, based on who they think the composer is. Therefore, it is expected that people who believe that AI can be creative will appreciate its music composed more than those who do not:

H1. There is a positive relationship between the belief that AI can be creative and the evaluation of its music.

Expectancy Violation Theory

Expectancy Violation Theory (EVT) explains individuals' reactions and responses to unexpected incidents or events in communication settings (Burgoon et al., 2016). The theory argues that when individuals see positive violations of expectations, that is, things are better than they had expected, they perceive the outcome as more favorable than if they had no expectations. A pleasant surprise disproportionately increases liking. Conversely, a negative violation within an interaction leads them to see the outcome as less favorable (Burgoon and Hale, 1988; Burgoon and Jones, 1976). An unpleasant surprise disproportionately decreases liking. A key element is therefore whether their expectation is met or not. In the context of this study, people will have an expectation about the quality of AI-composed music before listening to it. If the quality of the song is either higher or lower than their expectation, then we should expect EVT to predict an extra degree of approval or disapproval accordingly. If the quality of the song matches the preexisting expectation, EVT predicts a simple confirmation effect. If the quality does not match, it predicts outsized effects. Thus, the valence of the violation is a critical factor leading to a positive or negative violation.

While EVT began in the context of expectancies from human conversations and interactions (Hall, 1966), the theory has recently been applied to verbal and computer-mediated communication, including in human-computer interaction studies (Bonito et al., 1999; Edwards et al., 2016). For instance, a previous study conducted by the original creator of the theory (Burgoon et al., 2016) used it to explain how people perceive embodied agents that deviate from their social expectations both positively and negatively. In that case, the study found that positive violations increase positive perceptions, while negative violations did not. Similar applications have been conducted in research regarding people's expectation of whether an AI or a robot can write the news. A study found that a negative violation could lead to a more negative perception of machine authorship (Waddell, 2018). Similar to writing news and painting, composing music is presumed to have a certain expectancy when the consumer knows it was created by AI. Expectancy violations have been applied to music in multiple studies but are limited to human-composed music (Janata, 1995; Steinbeis et al., 2006). This presents a clear gap in the research with a consistent theoretical framework, but without empirical testing into AI music and reactions. This body of research and EVT offer a direct set of predictions about reactions to AI-composed music. This study will investigate how people react when the music composed by an AI is either unexpectedly better or worse than their expectations, suggesting:

H2. People who think the AI-composed music is better than expected will give higher ratings than people who think the music is within the range of expectation.

H3. People who think the AI-composed music is worse than expected will give lower ratings than people who think the music is within the range of expectation.

Music genre schema

When listening to AI musical numbers, machine heuristics are likely not the only factor that influences an evaluation. People's schema about the music genre would be another factor. A schema refers to the cognitive framework that is needed for perceiving, comprehending, and remembering (Brewer and Treyens, 1981). Its function is to decrease the complexity of information so that typical situations or familiar objects can be processed more efficiently with less mental effort (Harris and Sanborn, 2014; Kleider et al., 2008). Therefore, schema and bias are often used interchangeably (Dixon, 2006; Mikulincer and Shaver, 2001). If an individual has a certain bias toward a genre, then the bias is highly likely to influence the evaluation of a song of the genre.

Prior research found that people's music genre preference can be a decisive factor in their evaluations and judgments of a music piece (Istók et al., 2013). One experimental study found that different genres of website background music create dissimilar perceptions of the website owner (Yang and Li, 2013). Similarly, music genre preference was found to be used as a standard to judge others due to the stereotypes people hold about the genre of music (Lastinger, 2011). Both of the examples support that people's attitudes toward genres have influences on the evaluation of other people.

Classical music is often regarded as part of highbrow culture, less digestible to the general public (Deihl et al., 1983; Peterson and Kern, 1996), whereas electronic dance music (EDM) is seen as a more lowbrow and accessible genre by generating more active participation and immersion given its nature as dance music (Kruger et al., 2018). These dissimilar biases coming from the music genres can influence the music evaluation. However, no study that compares musical quality between different genres has been conducted, even between human-composed ones. Still, it is presumed that people's reactions to music would differ based on the genre.

H4. The evaluation of AI-created music will differ based on the genre.

H5. There will be an interaction effect between the genre of AI-created music and expectations toward it.

Methods

In order to test the hypotheses, a $2 \times 2 \times 2$ experiment was designed and conducted, where expectancy violations (expectancy confirmation vs violation), the valence of evaluation (positive vs negative), and genre of music (EDM vs classical) were different. The attitude toward the creation of AI was used as a covariate. The dependent variable was a subjective evaluation of the music.

Participants

Amazon Mechanical Turk (MTurk) was used to recruit participants. Individuals voluntarily joined the study after reading its purpose. Participants who failed an attention test (e.g. What is the genre of this music?) were excluded, leaving 299 participants from the 426 who were initially recruited (37.5 participants in each cell). The youngest participant was 19 years old, while the oldest was 73 years old ($M = 33.39$, $SD = 9.40$). In terms of gender, 61.9% of them identified as male, and 38.1% identified as female.

Procedures

Two AI-composed classical music pieces and two AI-composed EDM pieces were used for this study. The genre was chosen as a unit of analysis because most studies about music preference use it as a variable (Christenson and Peterson, 1988; Delsing et al., 2008; Ferrer et al., 2013; Schäfer and Sedlmeier, 2009). There are multiple differences between genres, such as the tempo, harmony, style, and structure. Therefore, having more than two genres without a clear difference would confuse the analysis of results. Instead, this study focused on two contrasting genres to make a more compelling case about the generalizability. Two pieces for each genre were chosen since a single song may not represent the whole genre. Those music pieces were chosen from a music library provided by a company called *Evoke Music*. The company approved the use of their music for this study.

To control for quality, a pilot test with participants recruited from MTurk ($n = 96$) was conducted to test whether the pieces in the two genres would be evaluated similarly. For this assessment, the fact that those songs are composed by AI was not told to participants. A t -test result confirmed that there was no significant quality between classical music ($M = 5.50$, $SD = 0.89$) and EDM ($M = 5.16$, $SD = 1.12$); $t(94) = -1.61$, $p = .111$. Also, a procedure to test whether the genres were distinct was followed. Participants were asked to choose the genre of the music they listened to. A chi-square test confirmed that the genres of the music used in the study, classical music and EDM, were distinct, not being confused with others; $\chi^2(1, N = 96) = 30.38$, $p < .001$.

After confirming that the stimuli were rated as equal in quality yet different in genre, an experiment using an online survey was conducted with a new set of participants. Participants were randomly assigned to one of the four songs (two songs per genre) and asked to listen to it. They were told whether their assigned piece was AI or human-composed before listening. Participants could not move on to the survey until the music ended. After listening to a given song, participants were asked to report their evaluation of a given musical piece, the level of expectancy violation with its valence, and their attitudes toward creative AI.

Measures

Expectancy violation scale. How much the music's quality deviates from their expectations was measured with a revised scale from an instrument used in a previous EVT study (Burgoon et al., 2016). This 7-point Likert-type scale (*Strongly disagree* to *Strongly agree*) measures both expectedness and evaluation. The three-item measurement of expectations shows how much it was violated (e.g. People would not be surprised to know that this music is composed by an AI), with higher scores indicating more expected outcomes ($\alpha = .82$). A five-item measurement was used for the valence of reactions to the music (e.g. Most people would find this AI-composed music enjoyable), with higher scores indicating more positive perceptions ($\alpha = .79$). Following the procedure used by Burgoon et al. (2016), the outcome of these measurements was used to distinguish the level and the valence of expectancy violations by conducting median splits.

Evaluation of music. The scales for evaluating music consisted of questions regarding their musical qualities. A 9-item scale for the assessment of musical quality was developed from the “Rubric for assessing general criteria in a composition assignment,” which is a scale used to measure multiple components in a musical composition ($\alpha = .90$) (Hickey, 1999). Components of the scale were aesthetic appeals (e.g. this AI-composed music piece presented a strong aesthetic appeal), creativity (e.g. the music piece included a very original musical idea), and craftsmanship (e.g. this AI-composed music had a clear beginning, middle, and end). Higher scores indicated a more positive evaluation of musical quality.

Attitudes toward creative AI. A scale that measured participants’ understanding of AI’s creativity was created and distributed. The 7-point Likert-type scale (*Strongly disagree* to *Strongly agree*) consisted of three statements: (a) I think AI can be creative on its own, (b) I believe AI can make something new by itself, and (c) Products developed by AI should be respected as creative works. Higher scores indicated more acceptance of creative AI ($\alpha = .86$).

Results

Simple linear regressions were conducted to test H1 that predicted a positive relationship between the perception of creative AI and the evaluation of AI-composed music. Overall, a significant and positive effect was found ($F[1, 297] = 92.71, p < .001$), with $r^2 = 23.8$ ($\beta = .49$). For classical music, a significant and positive effect was also found ($F[1, 171] = 58.14, p < .001$), with $r^2 = 25.4$ ($\beta = .50$). Finally, a significant and positive effect was also found for EDM ($F[1, 124] = 31.59, p < .001$), with $r^2 = 20.3$ ($\beta = .45$). An increase in positive attitudes toward creative AI, leading to higher ratings, was found regardless of the genre. Figure 1 shows a graph of a summary of the regression results. Based on the results, H1 was supported.

Two sets of two-way analysis of covariance (ANCOVA) were conducted to test H2, anticipating higher music evaluation ratings from the positive expectancy violation group (the music was above expectation) compared to the expectancy confirmation one (the music was as good as expected), and H3, expecting lower ratings from the negative expectancy violation group (the music was below expectation) than the expectancy confirmation one (the music was as bad as expected). Based on the valence of violation, the two ANCOVAs used the expectancy violation and the genre of music as independent variables. The subjects’ attitudes toward creative AI were used as covariates because the regression analysis above showed the relationship between the attitudes toward AI being creative and the evaluation of AI-composed music. The dependent variable was the evaluation of AI-composed music. Levene’s test was conducted to assess the equality of variances and rejected the homogeneity of variances for both the positive evaluation ($F[3, 130] = 0.58, p = .63$) and the negative evaluation ($F[3, 161] = 1.40, p = .25$), meaning that the ANCOVAs could be conducted. For the positive evaluation, the evaluation done by the expectancy violation group ($M = 5.91, SD = 0.67$) and the expectancy confirmation group ($M = 5.90, SD = 0.71$) showed no significant outcome ($F[1, 129] = 0.00, p = .99$). Therefore, H2 was rejected. In addition, no significant outcome was found from the genres ($F[1, 129] = 0.11, p = .74$), between classical music ($M = 5.90, SD = 0.71$) and EDM ($M = 5.91, SD = 0.67$), with no interaction effect ($F[1, 129] = 2.47, p = .12$). For the negative evaluation, on the other hand, a significant outcome was found ($F[1, 160] = 7.51,$

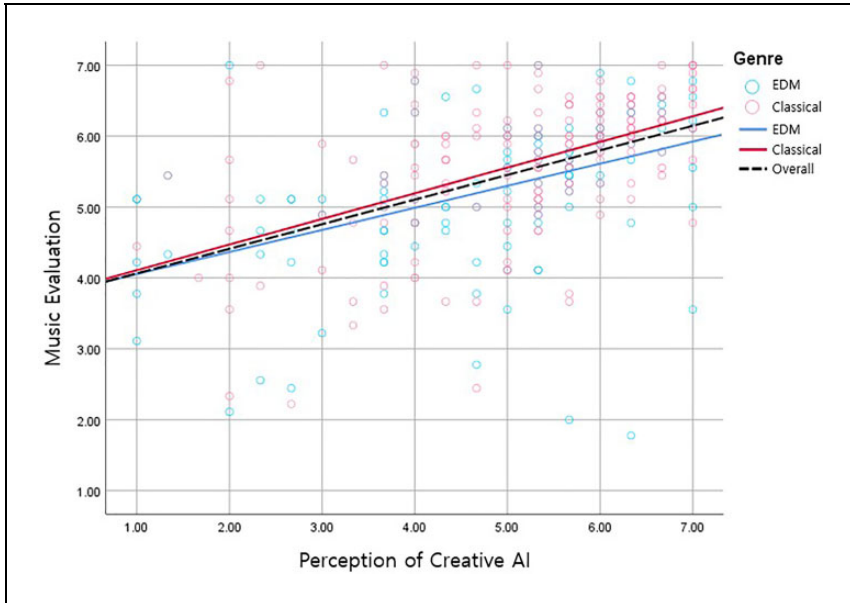


Figure 1. The regression analysis of the relationship between the perception of creative AI and the evaluation of AI-composed music.

$p = .007$, $\eta^2 = .05$) between the expectancy violation group ($M = 4.71$, $SD = 1.05$) and the expectancy confirmation group ($M = 5.39$, $SD = 1.07$). Therefore, H3 was supported. In addition, no significant outcome was found from the genres ($F[1, 160] = 1.26$, $p = .26$), between classical music ($M = 5.28$, $SD = 1.11$) and EDM ($M = 4.86$, $SD = 1.08$), with no interaction effect ($F[1, 160] = 0.45$, $p = .50$).

H4 focused on the influence of genres on music evaluation, and H5 on the effect of genres with expectancy violation and its valence. To answer those two questions, a three-way ANCOVA was conducted with the same covariate. A Levene's test rejected the homogeneity of variances of this ANCOVA ($F[7, 291] = 1.81$, $p = .09$). There was a significant difference between the expectancy violation ($M = 5.91$, $SD = 0.70$) and confirmation ($M = 5.08$, $SD = 1.11$) in terms of evaluating AI-composed music ($F[1, 290] = 4.17$, $p = .04$, $\eta^2 = .01$). However, no significant difference was found ($F[1, 290] = 0.78$, $p = .38$) between classical music ($M = 5.60$, $SD = 0.98$) and EDM ($M = 5.24$, $SD = 1.07$), which rejected H4.

While the genre had no interaction effect either with the expectation-violation ($F[1, 290] = 0.53$, $p = .38$) or its valence ($F[1, 290] = 1.40$, $p = .24$), there was a two-way interaction effect between the expectancy violation and its valence ($F[1, 290] = 5.57$, $p = .019$, $\eta^2 = .02$). Finally, there was a three-way interaction effect between the expectancy violation, its valence, and the genre ($F[1, 290] = 4.00$, $p = .05$, $\eta^2 = .01$). Therefore, H5 was supported. Figure 2 shows a graph of the ANCOVA results between the expectation-violation and its valence for classical music, and Figure 3 shows the ANCOVA results between the expectation-violation and its valence for EDM.



Figure 2. The ANCOVA results on music evaluation for classical music.

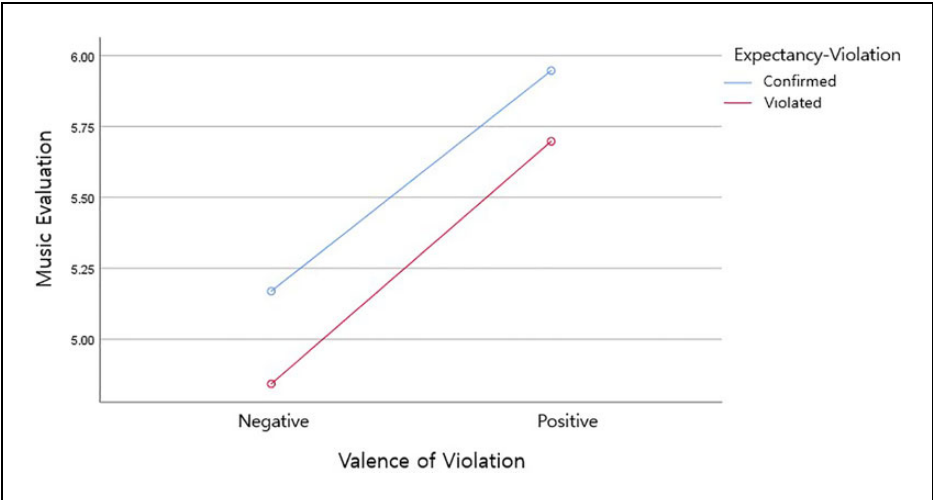


Figure 3. The ANCOVA results on music evaluation for EDM.

Discussion

The purpose of this study was to examine general reactions to AI-created music, as well as whether a predisposition would affect the results when the subject was positively or negatively surprised by the quality of the music. Strongly contrasting genres were also tested to see if the results would apply more broadly across musical types. The results suggest that accepting the creativity of AI is a prerequisite for a positive evaluation of its artistic merit. This positive correlation, regardless of the genre of music, shows that

appreciating the beauty of new things comes from an open-minded attitude, and that conversely, an unwillingness to accept AI products blocks appreciation. As mentioned earlier, unless there are commonly agreed upon objective criteria for a product to be deemed art, the definition of artistic performance by an AI relies on an individual's subjective criteria (Coeckelbergh, 2017). Even though people have similar beliefs and biases toward machines (Sundar and Kim, 2019), this does not mean that the strength of their attitudes is equal. By examining the perceptions of AI and music for the first time, this study finds that different predispositions toward creative AI can lead to a different appreciation of its products. Taken together with previous findings in visual design, this suggests that preexisting attitudes toward machines must be considered in HMC studies. Machine heuristics are speculated to influence the perception of AI, including its violations of expectations. The expectation is a different form of schema that comes from cumulated knowledge and experiences (Bonito et al., 1999), so the perception of expectancy violations from an AI's performance depends on their preexisting attitudes. Therefore, this study suggests that future EVT studies involving evaluations of AI performances should consider existing attitudes toward particular characteristics of AI in their design.

Also, a two-way interaction between the expectancy violation and its valence was found, which supports EVT. Participants with positive expectancy violations, who thought the given music was surprisingly better than their expectation about AI-composed music, reported higher ratings compared to the ones with positive expectancy confirmations, who thought music had a high quality, as anticipated. On the other hand, participants with negative expectancy violations, who thought the music was much worse than what they expected from AI-composed music, reported lower ratings compared to the ones with negative expectancy confirmations, who thought the music had low quality, as expected. This is a new extension of EVT into a new aspect of human-computer interaction. It is worth noting that while EVT was established as an interpersonal theory (Burgoon, 2015), the application of the theory expanded to non-interactive settings (Cohen, 2010; Rui and Stefanone, 2018), with room for more nuance there as well. This study has extended the use of EVT to now include evaluations of a machine's performance. However, the performance of machines cannot always be predicted. An overreaction to a machine's work might lead to an incorrect attitude toward the technology more broadly (Wakefield, 2016). Broader applications of EVT can help minimize unpleasant surprises from machines and lead to higher quality human-computer interactions.

While a previous EVT study about the AI perception tested the violation of expectations and its valence (Burgoon et al., 2016), this study further considered different characteristics of its performance now extended to music and perhaps to the way humans think about different genres. From the ANCOVA results, a three-way interaction effect between the expectancy violation, its valence, and the genre of music was found. While an interaction effect between the expectancy violation and the valence violation was found from classical music, there was no interaction effect for EDM. In other words, the genre matters. While the skewnesses of the expectation confirmation and the expectancy violation in the three-way ANCOVA graph about classical music were clearly different, the skewnesses in the EDM condition were similar. In other words, expectation-violation

matters in music evaluations for classical music but not for EDM. This was (ironically) *unexpected*, and suggests that some different processing of the two genres of AI-composed music lead to a different kind of expectancy violation that EVT has not directly considered before. This difference may be a theoretical contribution, if it is replicated by others. Why would it occur, and how can EVT be applied with new nuance to account for it? It has been pointed out that the intensity and effects of expectancy violation may not always be the same (Afifi and Metts, 1998). Why would this intensity be different here? It is speculative, but it may be that different levels of anthropomorphism in classical music and EDM had an influence on the evaluation of AI's performances. Electronic music has been regarded as "less human" compared to other types of music since the software is expected to be highly involved in its creation (Cookney, 2012; Eads, 2015). These dissimilar expectations toward genres based on their different level of anthropomorphic aspects may have influenced the evaluation of AI-composed music. In other words, AI-composed EDM may create less cognitive dissonance than AI-composed classical music. However, this study cannot confirm it since the expectation toward the musical anthropomorphism was not measured. Therefore, this study suggests future EVT studies about AI perceptions consider the level of anthropomorphism.

This study has a few limitations. First, attitudes toward a particular genre of music were not measured before the experiment. Therefore, it is not clear whether the genre preference has any effect on the evaluation of music quality. However, this could not account for the within-subject findings. Future studies should consider genre and attitudes toward it to test the speculations made here about the anthropomorphism of EDM music. In addition, because the AI-composed music pieces were retrieved from a single company, there may not be many variances of style within each genre. Songs created by other AIs may create different results. Using different AI composers may allow for better generalizability. Finally, future research might consider the subjectivity and taste of different genres. Qualitative assessments would potentially generate insights that this study might have missed.

The results of the study have clear implications for theory, namely the extension and nuance of EVT. The study broadens the applicability of EVT by using it to analyze the evaluation of AI performances in non-interactive settings. While EVT has been used to assess the communication satisfaction with a single AI agent, this study is the first to look at the perception of AI performances that does not involve a direct interaction. Application of EVT to AI perception is expected to be crucial since people assess AI based on whether their expectation is met or not (Burgoon et al., 2016). Previous human-computer interaction studies have focused on people's attitudes toward machines and their communication satisfaction with AI agents (Bartneck et al., 2007; Shank et al., 2019; Suen et al., 2019). However, people do not have identical perceptions of AI and the level of the perception varies. On the other hand, expectancy violations and confirmations can happen anytime, regardless of what beliefs or attitudes people have about machines. Therefore, this study suggests that future research should continue to test different contexts of human-AI interaction. Perhaps music, visual arts, and other media will work differently for EVT. Understanding which media and products lead to different outcomes

will add nuance to the theory, and allow researchers to better understand the “why” and “how” of EVT as it applies differently.

Also, the findings of this study have general implications. In order for the creative output of AI to be valued, people must first be persuaded that AI can be creatively autonomous. Attitudes toward new technologies are often stable within a generation, so it may be challenging to convince some age groups (Chung et al., 2010; Niehaves and Plattfaut, 2014). Also, companies should avoid negative expectancy violations. People are shown to devalue AI-composed music significantly when music was worse than expected (negative violation). On the other hand, people did not appreciate the music above their expectations significantly more compared to the music they thought as good as expected (positive violation). Therefore, fulfilling basic expectations is critical to avoid large-scale rejection of the technology. In order to do so, the understanding of people’s expectations about AI-powered products and services should be preceded, but how people perceive the creativity of AI has not yet been studied enough. Therefore, this study calls for further research about public perceptions of creative AI. The fact that AI can be creative refers to the possibility that it can provide new insights that have not been considered before. By rejecting the concept of creative AI, we may lose opportunities to have a broader perspective of the world. For instance, the emergence of creative AI brings collaboration between human and AI artists, instead of substituting the role of human artists (Feldman, 2017). Even though there have been many research studies about creating creative AI, there has been a lack of studies about how people see AI artists and their products. Knowing how people see AI as being creative is expected to aid the development of creative AI that people can accept. While this research only dealt with music composed by AI, it is hoped that its findings can contribute to future studies about creative AI.

Finally, this study is expected to broaden perspectives within human–computer interaction research more generally. Existing studies have focused on social interactions with machines, particularly looking at whether people interact with machines just as they do with human interlocutors. However, people interact with machines’ products, not just directly. Therefore, HMC can encompass artistic media such as music, dance, and visual art. For instance, music is used for clinical purposes (Alty et al., 1997; Cross, 2014; Hargreaves et al., 2005). By having an in-depth understanding of the role of AI-composed music in human–machine interactions, it may be possible to create music for specific purposes, such as developing an AI composer for music therapy. By investigating various media between human and machine interlocutors, the field of HMC is expected to have broader application.

Authors’ note

All authors have agreed to the submission, and the article is not currently being considered for publication by any other print or electronic journal.


Acknowledgements

Thanks to Evoke Music for allowing us to use their music for this study.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Joo-Wha Hong  <https://orcid.org/0000-0002-6555-3074>

References

- Adams RL (2017) 10 powerful examples of artificial intelligence in use today. *Forbes*. Available at: <https://www.forbes.com/sites/robertadams/2017/01/10/10-powerful-examples-of-artificial-intelligence-in-usetoday/#658cdafc420d> (accessed 19 July 2019).
- Afifi WA and Metts S (1998) Characteristics and consequences of expectation violations in close relationships. *Journal of Social and Personal Relationships* 15(3): 365–392.
- Alty JL, Rigas D and Vickers P (1997) Using music as a communication medium. In: *CHI'97 extended abstracts on human factors in computing systems, Atlanta, GA, 22–27 March*, pp. 30–31. New York: ACM. DOI: 10.1145/1120212.1120234.
- Baird B, Blevins D and Zahler N (1993) Artificial intelligence and music: implementing an interactive computer performer. *Computer Music Journal* 17(2): 73–79.
- Barreau P (2018) How AI could compose a personalized soundtrack to your life (Video file). *TED*. Available at: https://www.ted.com/talks/pierre_barreau_how_ai_could_compose_a_personalized_soundtrack_to_your_life (accessed 17 July 2019).
- Bartneck C, Suzuki T, Kanda T, et al. (2007) The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *AI & Society* 21(1–2): 217–230.
- Bharucha JJ and Olney KL (1989) Tonal cognition, artificial intelligence and neural nets. *Contemporary Music Review* 4(1): 341–356.
- Boden MA (1998) Creativity and artificial intelligence. *Artificial Intelligence* 103(1–2): 347–356.
- Bonito J, Burgoon JK and Bengtsson B (1999) The role of expectations in human-computer interaction. In: *Proceedings of the international ACM SIGGROUP conference on supporting group work, Phoenix, AZ, 14–17 November*, pp. 229–238. New York: ACM.
- Brewer W and Treyns J (1981) Role of schemata in memory for places. *Cognitive Psychology* 13(2): 207–230.
- Burgoon JK (2015) Expectancy violations theory. In: *The international encyclopedia of interpersonal communication*, pp.1–9. <https://onlinelibrary.wiley.com/doi/full/10.1002/9781118540190.wbeic102>
- Burgoon JK and Hale J (1988) Nonverbal expectancy violations: model elaboration and application to immediacy behaviors. *Communication Monographs* 55(1): 58–79.
- Burgoon JK and Jones SB (1976) Toward a theory of personal space expectations and their violations. *Human Communication Research* 2(2): 131–146.
- Burgoon JK, Bonito J, Lowry P, et al. (2016) Application of Expectancy Violations Theory to communication with and judgments about embodied agents during a decision-making task. *International Journal of Human-Computer Studies* 91: 24–36.
- Camurri A (1990) On the role of artificial intelligence in music research. *Journal of New Music Research* 19(2–3): 219–248.
- Chamberlain R, Mullin C, Scheerlinck B, et al. (2018) Putting the art in artificial: aesthetic responses to computer-generated art. *Psychology of Aesthetics, Creativity, and the Arts* 12(2): 177.
- Christenson PG and Peterson JB (1988) Genre and gender in the structure of music preferences. *Communication Research* 15(3): 282–301.

- Chung JE, Park N, Wang H, et al. (2010) Age differences in perceptions of online community participation among non-users: an extension of the Technology Acceptance Model. *Computers in Human Behavior* 26(6): 1674–1684.
- Coeckelbergh M (2017) Can machines create art? *Philosophy & Technology* 30(3): 285–303.
- Cohen EL (2010) Expectancy violations in relationships with friends and media figures. *Communication Research Reports* 27(2): 97–111.
- Cookney DJ (2012) Post-human pop: from simulation to assimilation. In: *Extremity and excess: proceedings of the 2011 University of Salford College of Arts and Social Sciences postgraduate research conference* (ed. Taylor E, Darlington J, Cookney D, et al.), Salford, 8–9 September, pp. 19–41. Salford: University of Salford Press.
- Cross I (2014) Music and communication in music psychology. *Psychology of Music* 42(6): 809–819.
- Curran NM, Sun J and Hong JW (2019) Anthropomorphizing AlphaGo: a content analysis of the framing of Google DeepMind’s AlphaGo in the Chinese and American press. *AI & Society*. DOI: 10.1007/s00146-019-00908-9.
- Deihl ER, Schneider MJ and Petress K (1983) Dimensions of music preference: a factor analytic study. *Popular Music & Society* 9(3): 41–49.
- Delsing MJ, Ter Bogt TF, Engels RC, et al. (2008) Adolescents’ music preferences and personality characteristics. *European Journal of Personality* 22(2): 109–130.
- Dixon TL (2006) Schemas as average conceptions: skin tone, television news exposure, and culpability judgement. *Journalism & Mass Communication Quarterly* 83(1): 131–149.
- Eads SJ (2015) Voice of the machine. *Senior Capstone Projects* 424. Available at: https://digital.window.vassar.edu/senior_capstone/424 (accessed 10 July 2019).
- Edwards C, Edwards A, Spence P, et al. (2016) Initial interaction expectations with robots: testing the human-to-human interaction script. *Communication Studies* 67(2): 227–238.
- Elgammal A, Liu B, Elhoseiny M, et al. (2017) Can: creative adversarial networks, generating “art” by learning about styles and deviating from style norms. arXiv:170607068.
- Feldman S (2017) Co-creation: human and AI collaboration in creative expression. In: *Electronic visualisation and the arts, London*, 11–13 July, pp. 422–429. Swindon: BCS Learning and Development Ltd..
- Ferrer R, Eerola T and Vuoskoski JK (2013) Enhancing genre-based measures of music preference by user-defined liking and social tags. *Psychology of Music* 41(4): 499–518.
- Furness D (2016) “Sunspring” is an absurd sci-fi short film written by AI, starring Thomas Middleditch. *Digital Trends*. Available at: <https://www.digitaltrends.com/cool-tech/sun-spring-ai-film-middleditch/> (accessed 18 July 2019).
- Gigerenzer G and Brighton H (2009) Homo heuristicus: why biased minds make better inferences. *Topics in Cognitive Science* 1(1): 107–143.
- Gigerenzer G and Gaissmaier W (2011) Heuristic decision making. *Annual Review of Psychology* 62: 451–482.
- Goode L (2018) AI made a movie—and the results are horrifyingly encouraging. *Wired*. Available at: <https://www.wired.com/story/ai-filmmaker-zone-out/> (accessed 18 July 2019).
- Guzman AL and Lewis SC (2020) Artificial intelligence and communication: a human–machine communication research agenda. *New Media & Society* 22(1): 70–86.
- Hall ET (1966) *The Hidden Dimension* (vol. 609). Garden City, NY: Doubleday.
- Hargreaves DJ, MacDonald R and Miell D (2005) How do people communicate using music. In: Miell D, MacDonald R and Hargreaves DJ (eds) *Musical Communication*. New York: Oxford University Press, pp. 1–26.
- Harris RJ and Sanborn FW (2014) *A Cognitive Psychology of Mass Communication*. 6th ed. New York: Routledge.

- Hewahi N, AlSaigal S and AlJanahi S (2019) Generation of music pieces using machine learning: long short-term memory neural networks approach. *Arab Journal of Basic and Applied Sciences* 26(1): 397–413.
- Hickey M (1999) Assessment Rubrics for Music Composition: Rubrics make evaluations concrete and objective, while providing students with detailed feedback and the skills to become sensitive music critics. *Music Educators Journal* 85(4): 26–52.
- Hong J and Curran NM (2019) Artificial intelligence, artists, and art: attitudes toward artwork produced by humans vs. artificial intelligence. *ACM Transactions on Multimedia Computing Communications, and Applications (TOMM)* 15(2S): 1–16.
- Hong J and Williams D (2019) Racism, responsibility and autonomy in HCI: testing perceptions of an AI agent. *Computers in Human Behaviors* 100: 79–84.
- Istók E, Brattico E, Jacobsen T, et al. (2013) “I love Rock ‘n’ Roll”—music genre preference modulates brain responses to music. *Biological Psychology* 92(2): 142–151.
- Jackson T (2017) Imitative identity, imitative art, and AI: artificial intelligence. *Mosaic: An Interdisciplinary Critical Journal* 50(2): 47–63.
- Janata P (1995) ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience* 7(2): 153–164.
- Jennings KE (2010) Developing creativity: artificial barriers in artificial intelligence. *Minds and Machines* 20(4): 489–501.
- Kleider H, Pezdek K, Goldinger S, et al. (2008) Schema-driven source misattribution errors: remembering the expected from a witnessed event. *Applied Cognitive Psychology* 22(1): 1–20.
- Kruger M, Viljoen A and Saayman M (2018) A behavioral intentions typology of attendees to an EDM festival in South Africa. *Journal of Convention & Event Tourism* 19(4–5): 374–398.
- Lastinger D (2011) The effect of background music on the perception of personality and demographics. *Journal of Music Therapy* 48(2): 208–225.
- Liu B and Wei L (2019) Machine authorship in situ: effect of news organization and news genre on news credibility. *Digital Journalism* 7(5): 635–657.
- Longuet-Higgins HC (1994) Artificial intelligence and musical cognition. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* 349(1689): 103–113.
- McCarthy J (2007) From here to human-level AI. *Artificial Intelligence* 171(18): 1174–1182.
- Marzano G and Novembre A (2017) Machines that dream: a new challenge in behavioral-basic robotics. *Procedia Computer Science* 104: 146–151.
- Meehan JR (1979) An artificial intelligence approach to tonal music theory. In: *Proceedings of the 19 annual conference 79* (eds Martin AL and Elshoff JL), Miami, FL, January, pp. 116–120. New York: ACM.
- Mikulincer M and Shaver PR (2001) Attachment theory and intergroup bias: evidence that priming the secure base schema attenuates negative reactions to out-groups. *Journal of Personality and Social Psychology* 81(1): 97–115.
- Niehaves B and Plattfaut R (2014) Internet adoption by the elderly: employing IS technology acceptance theories for understanding the age-related digital divide. *European Journal of Information Systems* 23(6): 708–726.
- Petersen MB (2012) Social welfare as small-scale help: evolutionary psychology and the deservingness heuristic. *American Journal of Political Science* 56(1): 1–16.
- Peterson RA and Kern RM (1996) Changing highbrow taste: from snob to omnivore. *American Sociological Review* 61(5): 900–907.
- Roads C (1980) Artificial intelligence and music. *Computer Music Journal* 4(2): 13–25.
- Roads C (1985) Research in music and artificial intelligence. *ACM Computing Surveys (CSUR)* 17(2): 163–190.

- Rui JR and Stefanone MA (2018) That tagging was annoying: an extension of expectancy violation theory to impression management on social network sites. *Computers in Human Behavior* 80: 49–58.
- Schäfer T and Sedlmeier P (2009) From the functions of music to music preference. *Psychology of Music* 37(3): 279–300.
- Shank DB, Graves C, Gott A, et al. (2019) Feeling our way to machine minds: people's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior* 98: 256–266.
- Shermer M (2017) Why artificial intelligence is not an existential threat. *Skeptic* 22(2): 29.
- Steinbeis N, Koelsch S and Sloboda JA (2006) The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience* 18(8): 1380–1393.
- Stere C and Trăușan-Matu Ș (2017) Generation of musical accompaniment for a poem, using artificial intelligence techniques. *Romanian Journal of Human–Computer Interaction* 10(3): 250–270. Available at: <http://search.proquest.com/docview/1989174700/> (accessed 18 July 2019).
- Sturm BLT, Iglesias M, Ben-Tal O, et al. (2019) Artificial Intelligence and music: open questions of copyright law and engineering praxis. *Arts* 8(3): 115.
- Suen HY, Chen MYC and Lu SH (2019) Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior* 98: 93–101.
- Sundar SS and Kim J (2019) Machine heuristic: when we trust computers more than humans with our personal information. In: *Proceedings of the 2019 CHI conference on human factors in computing systems* (eds Fitzpatrick G and Brewster S), Glasgow, 4–9 May, p. 538. New York: ACM.
- Turkle S (2005) *The Second Self: Computers and the Human Spirit*. Cambridge, MA: MIT Press.
- Waddell T (2018) A robot wrote this?: how perceived machine authorship affects news credibility. *Digital Journalism* 6(2): 236–255.
- Wakefield J (2016) Microsoft chatbot is taught to swear on Twitter. *BBC News*, 25 March. Available at: <https://www.bbc.com/news/technology-35890188>
- Yang Q and Li C (2013) Mozart or metallica, who makes you more attractive? A mediated moderation test of music, gender, personality, and attractiveness in cyberspace. *Computers in Human Behavior* 29(6): 2796–2804.
- Zulić H (2019) How AI can change/improve/influence music composition, performance and education: three case studies. *INSAM Journal of Contemporary Music, Art and Technology* 1(2): 100–114.

Author biographies

Joo-Wha Hong is a researcher in the Annenberg School for Communications at the University of Southern California. His work focuses on human–machine communication, particularly focusing on how people perceive artificial intelligence and interact with AI agents.

Qiyao Peng is a researcher in the Annenberg School for the University of Southern California. Her work focuses specifically on online communities and computer-mediated communication, especially in the field of new technology, video games, and artificial intelligence.

Dmitri Williams is an associate professor in the Annenberg School for Communications at the University of Southern California. His work focuses on communication, technology, and social capital using a systems lens. He has an emphasis on video games, analytics, and large-scale data.