What's Behind the Magic? Audiences Seek Artistic Value in Generative AI's Contributions to a Live Dance Performance

Jacqueline Elise Bruen Virginia Tech Blacksburg, Virginia, USA jacquelineb@vt.edu

Abstract

With the development of generative artificial intelligence (GenAI) tools to create art, stakeholders cannot come to an agreement on the value of these works. In this study we uncovered the mixed opinions surrounding art made by AI.

We developed two versions of a dance performance augmented by technology either with or without GenAI. For each version we informed audiences of the performance's development either before or after a survey on their perceptions of the performance. There were thirty-nine participants (13 males, 26 female) divided between the four performances. Results demonstrated that individuals were more inclined to attribute artistic merit to works made by GenAI when they were unaware of its use. We present this case study as a call to address the importance of utilizing the social context and the users' interpretations of GenAI in shaping a technical explanation, leading to a greater discussion that can bridge gaps in understanding.

ACM Reference Format:

Jacqueline Elise Bruen and Myounghoon Jeon. 2025. What's Behind the Magic? Audiences Seek Artistic Value in Generative AI's Contributions to a Live Dance Performance. In *Proceedings of Explainable AI for the Arts Workshop 2025 (XAIxArts 2025)*. ACM, New York, NY, USA, 6 pages. https: //doi.org/10.1145/nnnnnnnnnnn

1 Introduction

While the use of computers to help create art has been accepted by the general public, GenAI has brought this into question; GenAI can produce high-quality artistic media [3, 5, 9]. While research has been done on how GenAI might impact society [7, 12, 21], and individuals have written commentary on how it can either be a boon or bane to the art community [4, 8, 10, 11, 17], little research has been done to uncover the biases against GenAI art and how these biases shape viewers' reception of explanations of how GenAI art is made. The present study seeks to understand audience perception when we (RQ1) withhold information on how a technologically augmented artwork was made, and (RQ2) alter the type of technology used (GenAI versus traditional digital tools).

We developed two versions of a technologically augmented dance performance; one made using AI software and one without. We developed a Likert scale survey for audience members which included questions on the creative value of the performances' technology. To uncover potential biases, we withheld information on how the performances were made to half of our audience members until after the survey. Hence, in this 2x2 between-subjects design

XAIxArts 2025, Online 2025. ACM ISBN 978-x-xxxx-x/YYYY/MM https://doi.org/10.1145/nnnnnnnnnnn Myounghoon Jeon Virginia Tech Blacksburg, Virginia, USA myounghoonjeon@vt.edu

we held four different performances (AI/Tell Before, AI/Tell After, Non-AI/Tell Before, Non-AI/Tell After). We analyzed the results of the Likert scale [14] surveys using Mann-Whitney tests [16, 22] and found trends between different pairs of performance responses outlined in Table 1.

The present work has both practical and theoretical implications. From a practical perspective, artists can use our findings to inform how they make and present their work. Understanding how subjective works are presented to and received by audiences can inform fields that apply art and technology, such as advertising. From a theoretical perspective, our work investigates people's value judgments of art which incorporates different types of technology. These judgments can shape how we approach explaining AI, perhaps by emphasizing elucidations on specific areas of concern for those with biases. Our investigation highlights how an art viewer's beliefs can guide the XAI community towards explaining GenAI in an arts context.

2 Methods

2.1 Participants

For the two "Non-AI" performances, we used technology to alter the visuals and sound. Participants heard about technology either before the performance or after the performance and survey. Eight participants attended the Non-AI/Tell Before performance, (3 male, 5 female) with an average age of 23.75 (SD = 4.80). Ten participants attended the Non-AI/Tell After performance, (4 male, 6 female) with an average age of 25.30 (SD = 4.99).

In the "AI" performances, we incorporated AI technology. Participants were told about the technology in a similar fashion to the Non-AI performances. Twelve participants attended the AI/Tell Before performance, (4 male, 8 female) with an average age of 29 (SD = 7.11). Nine participants attended the AI/Tell After performance, (2 male, 7 female) with an average age of 23.89 (SD = 3.89).

2.2 Materials

We developed two versions of a performance with a professional dancer along with a survey. Figures 1 and 2 in Appendix A showcase the system architectures for each performance, and Figures 3 and 4 in the same appendix depict our development process. We used the Vernier Go Direct Respiration Belt and the Qualisys Track Manager (QTM) to obtain live data of the dancers' breathing and location during the performances. The Qualisys infrared cameras detected the reflective markers we positioned on the dancers' ankles. Live data were used to change the visuals on a large, mounted television screen behind the dancer and create a sonification for scene 2.

The dancer's choreography and the performance's structure were kept the same in the two versions. The performances differed in the use of technology. In the non-AI performances, the creative technological decisions were made by the technologist. In the AI performances, some of the creative portions of the technologist's work were made or supported by AI. The two versions of the 12 minute performance included three, four minute scenes. In scene 1, petals drifted through the sky with the dancer's respiratory rate. The scene was silent. Figure 1a in Appendix A depicts the AI imagery. In scene 2, waves moved with the dancer's location. Sonification of the dancer's position was played. Figure 1b in Appendix A depicts the AI imagery. In scene 3, stars brightened and multiplied with the dancer's breathing patterns. Figure 1c in Appendix A depicts the AI imagery. Music was chosen by the dancer.

In the Non-AI performance version, decisions on how to map data to visuals and sound were made by the technologist. The technologist sourced imagery from human-produced visual works, which were then made to dynamically change with the live data by layering multiple images and altering them conditionally according to data mappings designed by the technologist.

The data mapping for the petals in scene 1 increased the distance the petals moved from their last position and brightened the sky in the background when the dancer's breathing rate increased. The mapping for the waves in scene 2 moved the waves across three axes in accordance with the dancer's position in the room. The mapping for the stars in scene 3 increased the brightness of the stars as the dancer's breathing rate increased. Figure 6 in Appendix A shows the dancer during the performance.

The data mapping for the sonification was made by segmenting the stage into five equal sections along each of the three axes (x, y, and z). Each axis was mapped to a different musical scale. At any location where the dancer was, the audience would hear three notes (one note from each of the three axis scales) played in unison.

In the AI performance version, the technologist delegated the imagery and data mapping designs to AI (ChatGPT) [18], and in scene 2, three neural network models were built using three dance subroutines. Models were built using Fiebrink and Cook's Wek-inator [6]. Each neural network had four connected inputs, one hidden layer, and four nodes per hidden layer. The four inputs to train each model were the x and y coordinate sums for each of the four reflective markers attached to the dancer's ankles (two markers per ankle). The neural networks provided probabilities of each of these dance subroutines currently happening during the live performance. When asking the GenAI for data-mapping ideas, it produced many options. The technologist consulted the dancer to come to a decision. Our prompts to ChatGPT and the responses we used are in the Appendix B.1.

The survey (in the Appendix B.2) collected audiences' feedback on the performances. The survey consisted of 29 five point Likert scale questions for the "Tell Before" performances. For the two "Tell After" performances, participants could also add any data-mapping connections they may have noticed. The survey took 10 minutes.

To analyze our survey data, we performed Mann-Whitney tests. We compared the following audience sets: 1) AI/Tell Before vs. Non-AI/Tell Before, 2) Non-AI/Tell Before vs. Non-AI/Tell After, 3) AI/Tell Before vs. AI/Tell After, and 4) AI/Tell After vs. Non-AI/Tell After. Because we added six additional Likert scale questions to the surveys in the "Tell After" conditions, we only included the 29 baseline questions in our Mann-Whitney tests for all but the last comparative case (AI, Tell After vs. Non-AI, Tell After). In the "Tell After" condition, we included the six additional questions as dependent variables in our Mann-Whitney test.

2.3 Procedure

This research was approved by our institution's IRB. The procedure across all performances was the same except for the time we informed participants of technology. For the "Tell Before" performances, we informed participants before the performance started. For the "Tell After" performances, we informed them after they had taken the survey.

Audience members were led into the studio where they sat around the stage. For five minutes, audience members of the "Tell Before" performances heard technology's role in what they would see. The performance then began and lasted for 15 minutes. Afterwards, informed consent documents for the study were distributed; interested individuals who verbally consented to the researcher could scan a QR code to the survey, which took approximately 10 minutes. After the survey, audience members of the "Tell After" performances heard how technology was used.

3 Results

There was statistical significance (confidence interval = 95%) for a selection of questions after conducting Mann-Whitney tests for the four performance sets where one independent variable was held constant; these are shown in Table 1 in Appendix C.

Statistically significant survey response differences between the Non-AI performances indicated that those told about technology before were watching for relationships between the dynamic visuals and the dancer compared to those told after. Comparing the AI performance responses, those told before the performance more greatly agreed with the statement that the visuals appeared random. Those told about the AI after the survey responded significantly higher on statements related to the artistic merit of the work.

4 Discussion and Conclusion

For those who experienced the Non-AI condition, the audience told before rated higher on survey questions pertaining to thinking about patterns in the visuals. This audience group also thought the visuals were more distracting than the audience told after. These results were in line with work on the effects of learned value on attentional capture [1]. Once the audience was aware of the visuals, they became more salient to them.

For those who experienced the AI condition, the audience told after the survey ended rated significantly higher on survey questions pertaining to the artistic merit of the visuals. The audience told before the performance rated significantly higher on the survey question: "The projected visuals appeared random." Our results suggest that viewers' awareness of how a creative work is produced influences their perception of the work's artistic value. Participants in the AI/Tell After condition heard how AI was used, but they didn't hear how AI works, which diverges from the classical emphasis in XAI on explaining how AI algorithms work. As methods in developing the best GenAI tools are becoming somewhat of a Audiences Seek Artistic Value in Generative AI's Contributions

trade secret and are virtually impossible to reverse engineer the outputs of, the present study makes a case to the XAI community for a greater focus on transparency of the presence and power of GenAI in the arts and beyond. It could be valuable to design a future study integrating these two approaches towards XAI: knowledge of AI's presence and knowledge of AI's mechanics.

When comparing the results for the "Tell After" audiences, those who experienced the AI performance rated higher on survey questions related to their curiosity. Those who experienced the Non-AI performance rated higher on seeing the complementary nature of the sound. These results align with research on the behavior of GenAI and art interpretation. GenAI is known to make mistakes [13, 20] and prompters lack fine-grained control in driving the output [15, 19]. Those lacking context when viewing art can perceive any product made by an artist as intentional [2]. Thus, an audience can attribute meaning to artistic products of GenAI that were unintended by the prompter.

Participants who experienced the AI/Tell Before performance were more curious as to why certain artistic decisions were made. These participants may have been searching for artistic intention in the AI performance which could have been more apparent in the non-AI performance. We see from the higher audience ratings on the complementary nature of the sound in the non-AI performance that participants could draw relational conclusions between the elements of the non-AI performance more easily than the AI performance. This could have contributed to different judgments on the creative merits of the performances.

Finally, we utilized both GenAI (a large language model) and traditional AI (a manual machine learning tool to train and run simple neural network models) in developing our AI performances. These technologies vary greatly and may require different explainability approaches. The XAI community may benefit from isolating these AI tools to test new methodologies in XAI.

Though we took precautions in the present study, we acknowledge that there is always a possibility for bias. We did experience some technical issues while implementing the performance. In the present study we collected feedback from audience members to understand how their impressions of an artistic work changed when we (RQ1) withheld information regarding how the artwork was created, and (RQ2) used different types of technology. Using audience surveys, we found trends indicating that withholding information on how an artwork was produced can impact audience members' evaluation of the work.

References

- Brian A Anderson, Patryk A Laurent, and Steven Yantis. 2011. Learned value magnifies salience-based attentional capture. *PloS one* 6, 11 (2011), e27926.
- [2] David Bayles and Ted Orland. 2023. Art & fear: Observations on the perils (and rewards) of artmaking. Souvenir Press.
- [3] Z Epstein, A Hertzmann, L Herman, R Mahari, MR Frank, M Groh, H Schroeder, A Smith, M Akten, J Fjeld, et al. 2023. Art and the science of generative AI: A deeper dive (arXiv: 2306.04141). arXiv.
- [4] Ziv Epstein, Aaron Hertzmann, the Investigators of Human Creativity, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R. Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex "Sandy" Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith. 2023. Art and the science of generative AI. Science 380, 6650 (2023), 1110–1111. doi:10.1126/science.adh4451 arXiv:https://www.science.org/doi/pdf/10.1126/science.adh4451
- [5] Stefan Feuerriegel, Jochen Hartmann, Christian Janiesch, and Patrick Zschech. 2024. Generative ai. Business & Information Systems Engineering 66, 1 (2024), 111–126.

- [6] Rebecca Fiebrink and Perry R Cook. 2010. The Wekinator: a system for real-time, interactive machine learning in music. In Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)(Utrecht), Vol. 3. Citeseer, 2–1.
- [7] Fiona Fui-Hoon Nah, Ruilin Zheng, Jingyuan Cai, Keng Siau, and Langtao Chen. 2023. Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. 277–304 pages.
- [8] Danah Henriksen, Nicole Öster, Punya Mishra, and Lindsey McCaleb. 2024. Generative AI, Creativity, Culture, and the Future of Learning: a Conversation with Mairéad Pratschke. *TechTrends* (2024), 1–7.
- [9] IBM. [n.d.]. What is generative AI? https://www.ibm.com/think/topics/ generative-ai. Accessed: 1-4-2025.
- [10] Harry H. Jiang, Lauren Brown, Jessica Cheng, Mehtab Khan, Abhishek Gupta, Deja Workman, Alex Hanna, Johnathan Flowers, and Timnit Gebru. 2023. AI Art and its Impact on Artists. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (Montréal, QC, Canada) (AIES '23). Association for Computing Machinery, New York, NY, USA, 363–374. doi:10.1145/3600211.3604681
- [11] Caroline A Jones, Huma Gupta, and Matthew Ritchie. 2024. Visual artists, technological shock, and generative AI. (2024).
- [12] Kostas Karpouzis. 2024. Plato's Shadows in the Digital Cave: Controlling Cultural Bias in Generative AI. *Electronics* 13, 8 (2024), 1457.
- [13] Jeong Hyun Kim, Jungkeun Kim, Jooyoung Park, Changju Kim, Jihoon Jhang, and Brian King. 2025. When ChatGPT gives incorrect answers: the impact of inaccurate information by generative AI on tourism decision-making. *Journal of Travel Research* 64, 1 (2025), 51–73.
- [14] Rensis Likert. 1932. A technique for the measurement of attitudes. Archives of Psychology (1932).
- [15] Vivian Liu and Lydia B Chilton. 2022. Design guidelines for prompt engineering text-to-image generative models. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–23.
- [16] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.
- [17] L Mineo. 2023. If it wasn't created by a human artist, is it still art. The Harvard Gazette. Artikkeli. Luettavissa: https://news. harvard. edu/gazette/story/2023/08/isart-generated-byartificial-intelligence-real-art (2023).
- [18] OpenAI. 2024. ChatGPT (May 13 version) [Large language model]. https://chat. openai.com. Accessed: 2025-05-07.
- [19] Jonas Oppenlaender, Rhema Linder, and Johanna Silvennoinen. 2024. Prompting AI art: An investigation into the creative skill of prompt engineering. *International Journal of Human–Computer Interaction* (2024), 1–23.
- [20] James Prather, Brent N Reeves, Juho Leinonen, Stephen MacNeil, Arisoa S Randrianasolo, Brett A Becker, Bailey Kimmel, Jared Wright, and Ben Briggs. 2024. The widening gap: The benefits and harms of generative ai for novice programmers. In Proceedings of the 2024 ACM Conference on International Computing Education Research-Volume 1. 469–486.
- [21] Dirk HR Spennemann. 2024. Generative artificial intelligence, human agency and the future of cultural heritage. *Heritage* 7, 7 (2024), 3597.
- [22] Frank Wilcoxon. 1992. Individual comparisons by ranking methods. In Breakthroughs in statistics: Methodology and distribution. Springer, 196–202.

A Performance Development Imagery



Figure 1. System architecture of the Non-AI performance version.







Figure 3. The dancer rehearsing during an experiment of different projection strategies.



Figure 4. Outline of the choreography for Scene 2.

(a) AI Scene 1 imagery



(b) AI Scene 2 imagery



(c) AI Scene 3 imagery

Figure 5. Imagery for each scene of the AI performance version.

Figure 6. The dancer during the performance.

B Research Methods

B.1 ChatGPT Prompts and Used Responses

Prompt for Scene 1:

"If I had to make petals of a painting move according to the respiratory rate of a dancer how should I map this respiratory rate to the petals?"

Used Response for Scene 1:

"...the petals could gently rotate based on changes in the breathing depth and they could also rise and fall based on the breathing cycles."

Prompts for Scene 2:

- "If I had to make waves move according to a continuous stream of three probabilities representing how close a dancer's movements are to one of three movements, how should I map the three continuous probabilities to the wave movements?"
- "If I had to make melodious sound according to a continuous stream of three probabilities representing how close a dancer's movements are to one of three movements, how should I map the three continuous probabilities to sound? Please give me details on pitch and tempo."
- "What pentatonic, minor, and chromatic scales should I use?"
- "What octaves should I use for the C major pentatonic scale, the A minor scale, and the chromatic scale from D to A if I will be using them together? What instrument should I use to complement graceful dancing?"

Used Response for Scene 2:

"...assign each probability a scale or a set of pitches... Movement A can map to a pentatonic scale.. Movement B can use a minor scale, and Movement C can use a chromatic scale." ChatGPT further

Bruen and Jeon



Audiences Seek Artistic Value in Generative AI's Contributions

suggested we use a C major pentatonic scale, an A minor scale, and a chromatic scale from D to A for each movement respectively.

Prompt for Scene 3:

"If I had to make stars in a night sky change according to a continuous stream of data representing a dancer's respiratory rate how should I map this respiratory rate data stream to the stars?" *Used Responses for Scene 3*:

- "...assign one probability [of a classified dance move currently happening] to control the height or amplitude of the waves... another probability could influence the wave's direction or angle."
- "...you could also experiment with the mappings to see what creates the most engaging visual experience." We took this liberty by mapping the third dance move's resemblance probability to the horizontal theta cosine offset of the wave.
- "...make the stars move more rapidly across the sky, suggesting heightened energy for a higher respiration rate. Their movement could be subtle or more dynamic, depending on how you want to emphasize the breathing. For a slower respiratory rate, the stars can move more slowly or remain still for a tranquil effect."

B.2 Audience Survey

All survey questions were asked of all audience groups unless otherwise noted. All questions were answered on a Likert scale from (1 - "Not at all" to 5 - "Absolutely") unless otherwise noted.

Performance Related Questions - Visuals Throughout the performance, we projected various visuals as the dancer performed. Please answer the following questions based on the performance you just watched.

- I paid attention to the visuals.
- The projected visuals were pleasing.
- The projected visuals complemented the performance.
- The projected visuals enhanced the performance.
- The projected visuals made sense in the context of the whole performance.
- The projected visuals demonstrated artistic merit.
- The projected visuals were creative.
- The projected visuals seemed meaningful.
- The projected visuals were distracting.
- The projected visuals appeared random.
- I think I would feel the same way about the performance if it didn't have visuals.
- I tried to see if there was a pattern in the visuals.
- I wondered what caused the visuals to change. AFTER
- Which of the following do you feel to be the most accurate representation of the performance? (Question asked only to those in the "Tell After" condition) (Likert scale ranged from: 1 "The visuals mostly seemed to guide the dancer." to 5 "The dancer mostly seemed to guide the visuals.")
- If I knew exactly what the visuals represented, I think I would enjoy the performance more. (Question asked only to those in the "Tell After" condition)
- I thought about how the visuals were created.
- I wondered why the visuals were chosen.

• I tried to make a connection between what the dancer was doing and the visuals.

Performance Related Questions - Sound In part 2, when we showed the wave visuals, there were some sounds that you heard during the performance. Please answer the following questions about this audio you heard.

- I didn't pay attention to the sound in part 2.
- The sound in part 2 was pleasant.
- The sound in part 2 complemented the performance.
- The sound in part 2 enhanced the performance.
- The sound in part 2 fit with the other sounds in the performance.
- The sound in part 2 demonstrated artistic merit.
- The sound in part 2 was creative.
- The sound in part 2 was meaningful.
- The sound in part 2 was distracting.
- The sound in part 2 seemed random.
- I think I would feel the same way about the performance if it didn't have sound in part 2.
- I tried to see if there was a pattern in sound in part 2.
- I was curious about why the sound in part 2 changed. (Question asked only to those in the "Tell After" condition)
- Which of the following do you believe to be the most accurate representation of part 2 of the performance specifically? (Question asked only to those in the "Tell After" condition) (Likert scale ranged from: 1 "The sound mostly seemed to guide the dancer." to 5 "The dancer mostly seemed to guide the sound.")
- If I knew exactly why the sounds were chosen, I think I would enjoy the performance more. (Question asked only to those in the "Tell After" condition)
- I thought about how the sound in part 2 was created.
- I wondered why the sound in part 2 was chosen.
- I tried to make a connection between what the dancer was doing and the sound in part 2.
- I noticed a mapping between the dancer's movements and the sound in part 2. (Question asked only to those in the "Tell After" condition) (Response choices were "Yes," "No," "I don't recall.")
- Can you briefly describe what you think the mapping between the dancer's movements and the sound in part 2 might be? (Question asked only to those in the "Tell After" condition) (Answers were written responses)

C Quantitative Results

XAlxArts 2025, June 23, 2025, Online

Performance	Statistically Significant Survey Response Differences
Comparison	
Both told before,	• The projected visuals enhanced the performance ($Z = -2.667$, $p = .008$, the Non-AI performance rated higher, [AI
comparing tech-	version: $M = 2.00$, $SD = .739$; Non-AI version: $M = 3.25$, $SD = .886$]).
nology type	• The projected visuals were distracting ($Z = -2.025$, $p = .043$, the Non-AI performance rated higher, [AI version: $M =$
	2.00, $SD = 1.348$; Non-AI version: $M = 3.13$, $SD = .991$]).
	• The projected visuals appeared random ($Z = -2.130$, $p = .033$, the AI performance rated higher, [AI version: $M =$
	3.67, SD = 1.073; Non-AI version: $M = 2.62, SD = .916$]).
Both Non-AI,	• The projected visuals were distracting ($Z = -3.333$, $p = <.001$, the "Tell Before" performance rated higher, [Tell
comparing time	before: <i>M</i> = 3.13, <i>SD</i> = .991; Tell after: <i>M</i> = 1.30, <i>SD</i> = .483]).
told	• I tried to see if there was a pattern in the visuals ($Z = -2.789$, $p = .005$, the "Tell Before" performance rated higher,
	[Tell before: $M = 4.88$, $SD = .354$; Tell after: $M = 3.60$, $SD = 1.075$]).
	• I thought about how the visuals were created. ($Z = -3.180$, $p = .001$, the "Tell Before" performance rated higher,
	[Tell before: $M = 4.50$, $SD = .756$; Tell after: $M = 2.30$, $SD = 1.160$]).
	•]). I tried to make a connection between what the dancer was doing and the visuals. ($Z = -2.049$, $p = .040$, the "Tell
	Before" performance rated higher, [Tell before: $M = 4.88$, $SD = .354$; Tell after: $M = 3.90$, $SD = 1.370$]).
Both AI, compar-	• The projected visuals complemented the performance ($Z = -2.411$, $p = .016$, the "Tell After" performance rated
ing time told	higher, [Tell before: $M = 2.33$, $SD = .778$; Tell after: $M = 3.33$, $SD = 1.00$]).
	• The projected visuals enhanced the performance ($Z = -2.007$, $p = .045$, the "Tell After" performance rated higher,
	[Tell before: $M = 2.00$, $SD = .739$; Tell after: $M = 2.89$, $SD = 1.054$]).
	• The projected visuals demonstrated artistic merit ($Z = -2.501$, $p = .012$, the "Tell After" performance rated higher,
	[Tell before: $M = 2.83$, $SD = 1.030$; Tell after: $M = 4.11$, $SD = 1.054$]).
	• Being informed about the production of a piece of art would impact its monetary valueThe projected visuals
	appeared random ($Z = -2.698$, $p = .007$, the "Tell Before" performance rated higher, [Tell before: $M = 3.67$, $SD = 1.073$;
	Tell after: $M = 2.22$, $SD = .972$]).
Both told after,	• The projected visuals were distracting ($Z = -2.008$, $p = .045$, the AI performance rated higher, [AI version: $M = 2.56$,
comparing tech-	SD = 1.509; Non-AI version: $M = 1.30$, $SD = .483$]).
nology type	• I thought about how the visuals were created ($Z = -2.795$, $p = .005$, the AI performance rated higher, [AI version: M
	= 4.11, SD = 1.054; Non-AI version: M = 2.30, SD = 1.160]).
	• I wondered why the visuals were chosen ($Z = -2.162$, $p = .031$, the AI performance rated higher, [AI version: $M =$
	4.44, $SD = .726$; Non-AI version: $M = 3.50$, $SD = .972$]).
	• The sound in part 2 complemented the performance ($Z = -2.061$, $p = .039$, the Non-AI performance rated higher,
	[AI version: $M = 3.56$, $SD = .882$; Non-AI version: $M = 4.30$, $SD = .483$]).
	• The sound in part 2 enhanced the performance ($Z = -2.253$, $p = .024$, the Non-AI performance rated higher, [AI
	version: $M = 3.78$, $SD = .667$; Non-AI version: $M = 4.50$, $SD = .527$]).
	• The sound in part 2 fit with the other sounds in the performance ($Z = -1.973$, $p = .048$, the Non-AI performance
	rated higher, [AI version: $M = 2.44$, $SD = .882$; Non-AI version: $M = 3.50$, $SD = 1.269$]).
	• I wondered what caused the visuals to change ($Z = -2.201$, $p = .028$, the AI performance rated higher, [AI version:
	M = 4.33, $SD = .707$; Non-AI version: $M = 3.20$, $SD = 1.135$]).

 Table 1: Statistically significant survey responses rendered from Mann-Whitney tests