JESK J Ergon Soc Korea 2024; 43(6): 473-484
https://dx.doi.org/10.5143/JESK.2024.43.6.473
https://jesk.or.kr eISSN:2093-8462

A User Study on Age-related Variances in the Perceived Emotional Alignment of Generative AIs

Yoon Kyung Lee, Yong-Ha Park, Sowon Hahn

Seoul National University, Department of Psychology, Seoul, 08826

Corresponding Author
Yoon Kyung Lee
Seoul National University, Department of Psychology, Seoul, 08826
Email : yoonlee78@snu.ac.kr

Received : April 16, 2024
Revised : August 02, 2024
Accepted : December 02, 2024

Copyright@2024 by Ergonomics Society of Korea. All right reserved.

Objective: This study aims to assess whether different age groups of generative AI users perceive AI's emotional intelligence differently, especially in its ability to align with users' emotions.

Background: Generative AIs have been rapidly implemented into users' lives, yet the potential and challenges of these models remain uncertain. We investigated how well these models understand users' emotions, specifically by age-related characteristics, which remain under exploration.

Method: Participants from various age groups (20 to 60, N=283) participated in the study and evaluated generative AI artwork generated from specific prompt instructions to reflect the writer's emotions (Emotion-focused; EF). Then, these artworks were compared to the ones generated from prompt instructions that were to reflect only the facts about the story (Information-focused; IF). Then, participants rated 240 images based on how they align with the user's emotions reflected in the story.

Results: Older adults (40s, 50s, and older) perceived AI more favorably in their alignment with user emotion than younger adults (20s and 30s). EF prompts outperformed IF prompts with significantly higher emotional alignment scores.

Conclusion: Contrary to popular belief, older adults were more favorable to AI-generated works than younger adults. Both younger and older adults preferred AI-generated work that reflects human emotions over those do not.

Application: Our findings highlight the importance of enhancing AI's emotional intelligence in increasing various demographic user engagement, especially in older adults. With increasing markets targeting senior users, these generative AIs can potentially increase accessibility and adoption rates in creative activities and businesses targeting well-being, healthcare, or education.

Keywords: Generative AI, AI in user experience, Emotional alignment, User study, Aging and AI

1. Introduction

The introduction of the generative adversarial network (GAN; Goodfellow et al., 2014) framework marked a pivotal moment in the development of generative AI, drawing attention from various domains. Generative AIs have showcased surprising capabilities, from human-like conversations to creating real-life images with simple natural language instructions or prompts (Brown et al., 2020; Ramesh, et al., 2021).

AI-generated images are becoming increasingly popular due to platforms like DALL-E, Midjourney, and open sources like Stable Diffusion (Ramesh et al., 2021; Rombach et al., 2022). These models generate images with a simple instruction prompt, from graphic designs to abstract art pieces (Slack, 2024). For example, users can generate an image with simple directions like 'I am confused and frustrated when using Zoom for the first time in the workspace' (which is shown in Figure 1).

Table 1 shows state-of-the-art generative AIs encompassing different modalities (text-to-text to text-to-image models). Many previous studies utilizing AIs have focused on emotion recognition, employing machine learning or deep learning as classifiers of emotions from a set of emotion categories and determining users' experiences based on facial expressions.

Table 1. Generative AIs released to date

Year	Company	Application	Modality	Model (s)	Reference
2022	Midjourney Inc.	Midjourney	TI	Midjourney	(Rombach et al., 2022)
2022	OpenAI	DALL-E	TI	DALL-E 2, 3	(Ramesh et al., 2021)
2022	Stability AI	Stable Diffusion	TI, II	Stable Diffusion	(Rombach et al., 2022)
2023	Microsoft	Microsoft Copilot	TI, TT	GPT-4, GPT-4-Vision	(Cambon et al., 2023)
2023	OpenAI	ChatGPT	TI, TT	GPT-3.5, GPT-4, GPT-4-Vision	(Brown et al., 2020)

Note: TI = Text to Image, TT = Text to Text, II = Image to Image

Past studies using generative AIs, such as those using GAN (Goodfellow et al., 2014) or CLIP (Radford et al., 2021) for style transfer of image types, have not concentrated on how the models interpret emotions human users describe. In specific, recent generative models, mostly text-to-text, focus on classifying emotions from the text (Table 1). However, evaluations of these generative models' accuracy, specifically those utilizing LLM, have been limited to their understanding of prompt instructions (Liu and Chilton, 2022), not reflecting their natural use in daily life. Thus, studies are scarce on how descriptions of emotions by human users are represented across different modalities, such as images.

Although many people increasingly use these generative models worldwide due to their simplicity and efficiency, such as winning the art contest (Roose, 2022), these models are still under development. We still don't know the potential and limitations of these models when different users use them. More importantly, it is unknown whether these models clearly understand human instruction and generate output that aligns with human intention – also known as 'AI alignment' (Gabriel, 2020). Specifically, it is important to know how generative AI can understand various contexts that represent users' everyday experiences and emotions.

Therefore, testing these in various contexts is important, focusing on human user emotion. Specifically, we define this as 'emotional alignment'—whether the AI model accurately understands the meaning of emotional words in the system instruction and accurately represents them in the output in a way the user intended. This is different from user satisfaction, as it does not specifically address whether the output is based on the 'understanding' of instruction provided by the human (whether it aligns it with the instruction or not, an output can still be perceived as satisfactory).

We, therefore, aim to explore how these models can understand human narrative without explicitly coding specific image components—which requires extensive knowledge such as hue, RGB code, camera angle, etc. The current study uses actual text from people who wrote about their daily lives and evaluates the generated images based on this human input on emotional alignment.

2. Method

2.1 Materials

We developed two main components: prompt instruction for the text-to-text generative model and output images from the text-to-image generative model. We used ChatGPT (https://openai.com/chatgpt/), a web interface that allows users to converse in natural language with GPT-series models for the text-to-text model. We used Stable Diffusion (stability.ai/stable-image) and DALL-E 2 (https://openai.com/index/dall-e-2/) text-to-image models to generate images.

2.1.1 Prompt development

Figure 1 illustrates the overall procedure of creating text-to-text prompts and text-to-image prompts. We created four unique images per diary entry by carefully selecting and capturing the essence of the events and emotions described under two prompt conditions: 'Emotion-focused (EF)' and 'Information-focused (IF)'. We then asked participants to rate whether they reflected the user's intended emotion from the text in the generated image.

For text input, we used human-written daily experiences selected from an available public dataset released by Lee et al. (2021). The data consists of text and the primary emotion expressed in the diary entry. Emotion is categorized based on Ekman's (1992) model, which includes happiness, sadness, anger, disgust, surprise, fear, and additional emotions such as calmness and neutrality. The topics of the dataset are composed of various domains, including relationships, school, work, health, personal stress or concerns, and the COVID-19 pandemic. Each text was composed of 250 characters (Korean).

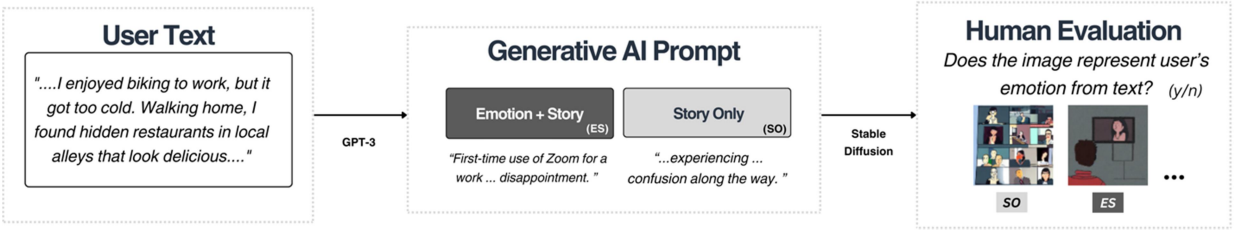

Figure 1. Overall procedure of the user study

Figure 2 illustrates the process of summarizing the user text (because it was too long, we had to shorten it and summarize it) and generating a text-to-image prompt format. Our original input was typically lengthy and contained, on average, 8.20 English sentences and 147.83 English words (8.23 Korean and 91.40 Korean words), making it usually too verbose for the model to process efficiently (extract only essential information and emotion keywords).

We selected 15 text inputs per survey with different categories of emotions reported by the original author (two diary entries of calmness, one neutral, six happiness, one anger, two sadness, one fear, one disgust, and one surprise). Each diary entry contains multiple emotional words, often making pinpointing one specific emotion representing the entire narrative difficult. People often struggle to name specific emotions, especially when experiencing multiple emotions simultaneously. To avoid such complexity, we extracted and summarized the sentences that contain the most representative emotions as labeled by the original author.

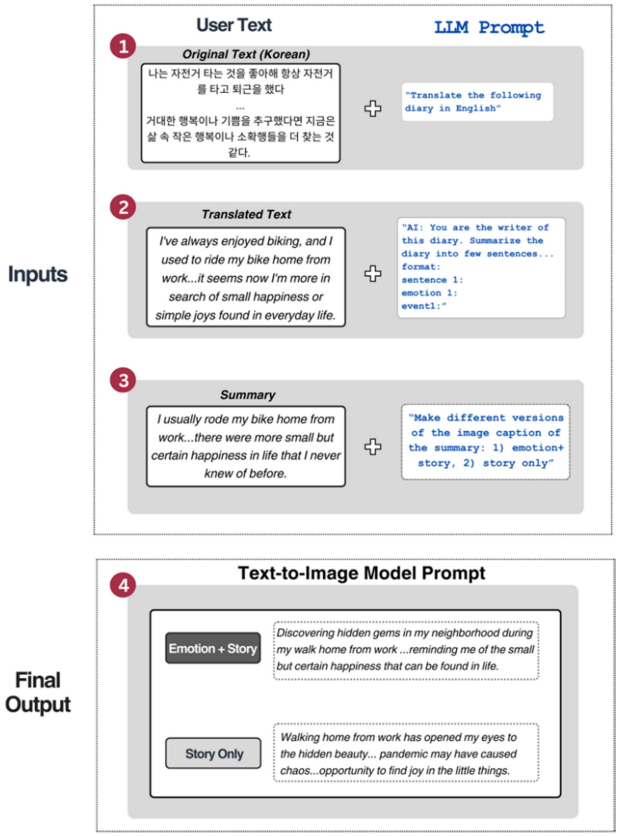

Figure 2. Process of tailoring prompts for the text-to-image model from human-written natural input. In this figure, the 'Original Text (Korean' input for the first step reads: "I've always enjoyed biking, and I used to ride my bike home from work … It seems now I'm more in search of small happiness or simple joys found in everyday life".

Cultural factors also play a significant role in expressing and categorizing emotions (Markus and Kitayama, 2014; Kitayama et al., 2006). To ensure the generated image accurately reflects the original writer's intended emotional expression, we referenced the main emotion based on the category the writer labeled. This approach led to reporting more complex social emotions, like calmness and neutrality.

We referred to neutral affect as "feeling indifferent, nothing in particular, and a lack of preference one way" (Gasper et al., 2019). Due to its neutral aspect, both positive and negative emotions can coexist. This is distinct from calmness, *차분함(chah-boon-ham)* or *평온함(pyung-on-ham)*, which some people may report as neutral but has a more positive sentiment than *덤덤함 (duhm-duhm-ham)*, as it is more related to being in a state of composure and peace.

- Step 1: Translation of Original Text: The original text from the emotion diary entry, written in Korean, was translated into English. This step aimed to ensure accurate natural language input processing and increase generalizability, as most models are primarily trained in English. The prompt instruction was: "Translate the following diary in English". We further ensured that the translated text reflected the main story of the diary entry input. Although some information may be translated in a more suitable way for the English context rather than the original Korean context, our purpose was to reflect the main story and emotion of the diary entry rather than the specific wording or syntax. We confirmed that most of the translated content accurately reflects the original content.

- Step 2: Summarization of Translated Text: We created a summary of the original text input because the original text was too long to be used as a prompt for image generation, and interpreting a single image output from a long prompt that contains too many events may be confusing. To avoid confusion, we instructed the model to summarize the diary into two sentences, one describing the main emotion and the other describing the main event. The prompt instruction was: "AI: You are the writer of this diary. Summarize the diary into two sentences. Format: sentence 1: emotion 1; sentence 1: event 1". If the diary entry included multiple cases, we added a format for the next sentence (e.g., sentence 2, emotion and event for sentence 2, but only if necessary).

- Step 3: Image Caption Generation: We used the summary from Step 2 and generated different styles of image captions based on the prompt conditions. We instructed the model to create multiple captions based on the summary, with variations emphasizing either "emotion + story" (Emotion-focused) or "story only" (Information-focused). The prompt was "make different versions of the image caption of the summary: 1) emotion + story, 2) story only".

- Step 4: Final text-to-image model prompt: The final step involved formulating prompts for a text-to-image model based on the created image captions. We refined the image captions to ensure the prompt generated the intended output. We also removed ambiguous expressions, unnecessary prepositions, and verbose phrases. This refinement aimed to produce high-quality images accurately representing the main story and the writers' emotions expressed in the diary entry. We finalized two prompts with two conditions: EF and IF prompt.

  - Emotion-focused prompt (EF prompt) included both the main emotions and the events.
  - Information-focused prompt (IF prompt): the prompt focused solely on describing the depiction of the main event (and trying to be more objective), rather than describing the inner experience of the writer.

2.1.2 AI-generated images

Sixteen images were prepared for each diary entry, spanning four different prompts. We further enhanced the prompts by specifying art styles ('abstract', 'figurative', 'watercolor', 'oil painting', 'monochromatic', 'and pop art'), and examples are shown in Table 2.

From the text-to-image generation level, we used various art styles because these can influence the perception of emotion expressed and are highly sensitive to personal taste. Different art styles combined 16 artworks applied to each text input in the study: 8 abstract and 8 figurative combinations or 4 watercolor, 4 oil paintings, 4 monochromatic, and 4 pop art combinations.

Table 2. Different styles of images generated by AI from the same text input

Prompt: *"Discovering hidden gems in my neighborhood during my walk home from work has been a delightful surprise, reminding me of the small but certain happiness that can be found in life".*

Abstract	Figurative	Watercolor	Oil Painting	Monochromatic	Pop Art

2.2 Research design

The independent variable was prompt conditions: *Emotion-focused* (prompts to generate 'emotion + story' from human text) and Information-focused (prompts to generate only the information or main event).

The dependent variable was the '(emotional) alignment score', assessed on a binary scale (1 for emotionally aligned, 0 for not emotionally aligned). Participants were asked to evaluate each image's alignment with the corresponding text prompt and whether they accurately represented the main event described in the prompt. The use of a binary rating scale was to ease the evaluation process for participants as they had to review many numbers of images (15 text inputs × 16 images to evaluate per text input = 240 images to evaluate). The order of images generated by the same text entry was randomly displayed.

Table 3 presents example images generated from different prompt conditions. For each text-images set, participants were asked to rate 8 images from the EF prompt and 8 from the IF prompt, totaling 16 images per text. We prepared 15 text inputs, which was an adequate number of questions for the participants to finish.

Table 3. Example image output per prompt condition

Summary of event	*"Because of the pandemic, had a video conference with a buyer overseas, and struggled".*	
Emotion-focused	*"Navigating the challenges of video conferencing during the pandemic, experiencing frustration and confusion along the way".*	
Information-focused	*"First-time use of Zoom for a work video conference during the pandemic led to unexpected difficulties and disappointment, highlighting the challenges of virtual communication".*	

2.3 Experiment procedure

Participants were informed about the study and that they would evaluate the various images generated by AI. The evaluation was done via the Qualtrics survey platform (www.qualtrics.com), a popular survey platform.

On the crowdsourcing platform, participants were shown a brief description of the study and informed consent, and they were to participate voluntarily. The platform provided the external link to the survey, and on the first page, a brief description of the task was shown to explain clearly what they were about to do. Then, they were asked to check 'confirm' before proceeding. Participants were informed that they could start, pause, or end the survey at their convenience, conforming to the nature of the online study.

The task is to evaluate whether the original writers' emotion is appropriately reflected in the image generated, but at the same time, the main story of the emotion is well reflected. This study was approved by the institution's IRB (No.2304/004-012).

3. Results

A total of 283 online participants completed the evaluation through the crowdsourcing platform *DeepNatural* (https://deepnatural.ai/) with an average age of 41.63 (SD = 8.39, N_{male} = 59). The minimum reported age was 22, and the maximum was 61, showing a diverse age range. The median age was 41.

3.1 Mean differences in AI emotional alignment between prompt conditions

Table 4 shows the mean differences between the two conditions. We calculated a total score by summing all alignment scores across 120 questions per condition (15 text inputs × 8 images per condition). Analysis revealed a significant difference in mean alignment scores between the images from the EF prompt (M = 56.82, SD = 22.65) and images from the IF prompt (M = 52.08, SD = 22.66) conditions ($t(282)$ = 10.13, p < .001).

Table 4. Mean AI emotional alignment score by prompt condition

	Emotion-focused	Information-focused
Mean (± SD)	56.82 (±22.65)	52.08 (±22.66)
Mean score in Percentage	47%	43%

Possible max score = 120 per prompt condition, SD = Standard Deviation

3.2 Mean differences of AI emotional alignment by age group

Table 5 and Figure 3 show mean differences in emotional alignment scores among age groups and between prompt conditions.

Table 5. Mean AI emotional alignment scores by age group

Age groups	N	Mean (SD)	
		Emotion-focused	Information-focused
20s	15	45.87 (±19.05)	41.60 (±18.27)
30s	99	47.87 (±19.15)	42.33 (±18.73)
40s	104	60.24 (±22.26)	56.66 (±22.14)
50s and older	65	67.49 (±23.14)	62.03 (±23.62)

All participants in each age group rated higher perceived emotional alignment in the EF prompt condition than in the IF prompt condition. Participants in their 50s and older had the highest scores in both conditions, and participants in their 40s and 50s and older showed the most substantial differences between conditions.

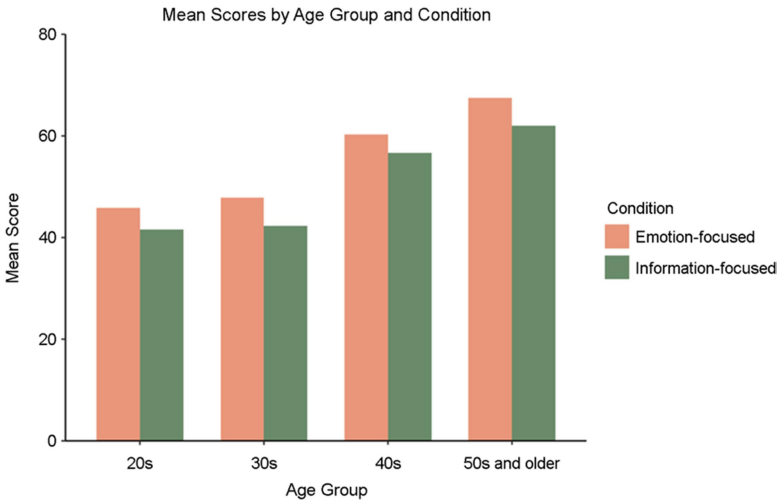

Figure 3. Mean score difference by prompt conditions and age group

3.2.1 ANOVA result

A two-way repeated measures ANOVA with 2 within (prompt conditions) × 4 between (age groups) mixed design was used for analysis. This examined the effects of prompt condition (EF and IF) and age group (4 groups) on emotional alignment (Table 6). The statistical significance criteria were set at $p < .05$. Assumptions for ANOVA were met, with only minor deviations from normality in the residuals as indicated by the Shapiro-Wilk test ($W = 0.99$, $p < .05$). These deviations were not problematic due to the larger sample size ($N = 283$) and diverse demographic characteristics. Levene's test for homogeneity of variance confirmed equal variances across groups ($F(7,558)=1.0867$, $p = .370$). The analysis was performed using R studio.

There was a significant main effect of group on total scores, $F(3, 558) = 28.03$, $p < .001$, $\eta^2 = .13$, indicating substantial differences among age groups. The effect of the condition was also significant, $F(1, 558) = 7.03$, $p = .008$, $\eta^2 = .01$, showing that the condition influenced total scores. However, the interaction effect between age group and condition was not significant ($F(3, 558) = 0.09$, $p = .966$, $\eta^2 < .001$).

Table 6. Two-way ANOVA result

Source of variation	SS	df	MS	*F*	*p*-value	η^2
Age group	37,905	3	12,635	28.026	< .001***	0.129
Prompt condition	3,168	1	3,168	7.026	.008**	0.011
Age group × Prompt condition	120	3	40	0.089	.966	0.0004
Residuals	251,569	558	451			

*$p < .05$. **$p < .01$. ***$p < .001$

3.2.2 Post Hoc analysis

Figure 4 and Table 7 present the results of the post hoc analysis, with significant p-values annotated. A Tukey's HSD test was

performed following the ANOVA to identify specific group and condition differences. Perceived emotional alignment was significantly higher for the EF prompt than the IF prompt in the older adult group. Specifically, participants in their 40s perceived higher emotional alignment than those in their 20s (Difference = 14.72, p = .002) and those in their 30s (Difference = 13.35, p < .001).

Older adults in their 50-60s showed significantly higher perceived emotional alignment compared to those in their 20s (Difference = 21.03, p < .001), 30s (Difference = 19.66, p < .001), and 40s (Difference = 6.31, p = .04).

There was a significant difference between EF and IF prompt conditions (Difference = -4.73, p = .008).

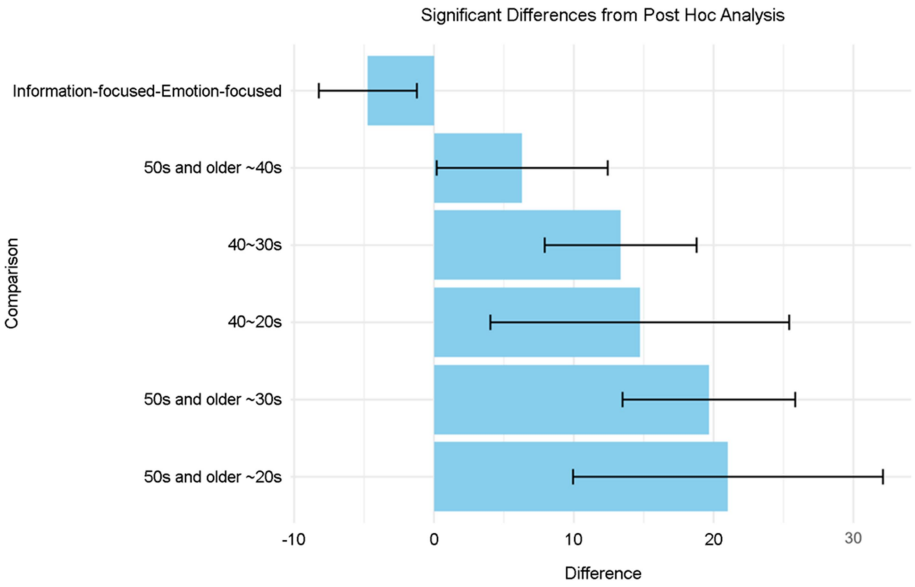

Figure 4. Significant differences from Post Hoc analysis for age group and condition comparisons

Table 7. Post Hoc test results for all group and condition comparisons

Comparison		Mean difference (M)	SE	T-value	CI	Adjusted p-value
Group	30~20s	1.37	5.47	0.25	[-9.35, 12.09]	.988
	40~20s	14.72	5.45	2.70	[4.03, 25.40]	.002**
	50s and older ~20s	21.03	5.65	3.72	[9.95, 32.11]	< .001***
	40~30s	13.35	2.77	4.82	[7.92, 18.78]	< .001***
	50s and older ~30s	19.66	3.15	6.24	[13.48, 25.84]	< .001***
	50s and older ~40s	6.31	3.12	2.02	[0.19, 12.43]	.040*
Condition	IF-EF prompt	-4.73	1.79	-2.65	[-8.24, -1.23]	.008

M = Mean difference, SE = Standard Error, CI = 95% Confidence Interval, *p < .05, **p < .01, ***p < .001

4. Discussions

Our study highlights the importance of emotional alignment in evaluating AI-generated artworks, especially across different age groups. It challenges some commonly held beliefs about older people and technology usage and suggests that emotional content could help bridge the gap between generations when it comes to tech adoption. This highlights the importance of emotional engagement in designing and deploying technology, particularly for older adults' needs and preferences.

The evidence suggests that as individuals age, their emotional response to AI-generated artwork undergoes a significant shift. Previous research has often highlighted older adults' challenges with adopting new technology, frequently attributed to perceived complexity or irrelevance (Czaja et al., 2006). However, our results suggest a different dimension: older adults may exhibit a higher receptiveness or appreciation for technology outputs, particularly when these outputs resonate emotionally. Specifically, we found that older age groups, particularly those in their 50s and older, consistently rated images higher than younger participants. This pattern was evident with significant differences observed between these age groups, indicating a stronger inclination or greater capacity among older adults to connect with or appreciate the AI-generated artwork, irrespective of the prompt type. This insight aligns with theories suggesting that emotional content becomes more salient and valued with age (Carstensen et al., 2003), potentially influencing how technology is perceived and evaluated.

We also found some common patterns in these generative models' limitations in portraying human emotion despite their impressive and efficient proficiency in processing natural language instructions. For example, AI artworks often overemphasized negative emotions rather than the degree to which they were depicted in the text. This might have influenced the users to rate some misaligned artworks with low scores (usually the ones visualizing 'sad' emotion into too depressing abstract artwork). Sometimes, too much abstract visualization of concrete content (e.g., a desk or a cup of coffee) would have influenced a lower alignment score. Future studies should consider whether the presence or absence of a certain context or object in AI-generated artwork affects AI alignment. Nonetheless, we found that the more the prompt was focused on the user's emotion (analogous to the user's narrative) than a simple description of listing the facts, the more people preferred the output generated from the former type of instruction.

In addition, the structure of the AI prompt, whether it focuses on emotion or information from the input, significantly influences the acceptance and appreciation of generative AIs. No significant interaction between age group and condition was found, suggesting their combined influence doesn't further modulate this effect. This underscores the need for more research about the potential influence of other factors, such as individual differences or level of expertise with generative AIs.

Lastly, from the perspective of AI's ethical and societal implications, the process of AI and humans is not the same, and people should be aware of anthropomorphizing AI's ability to reflect human emotions (Epstein et al., 2020). One other point that should be considered in the development of AI models is demographic diversity, which, in the current study, we delved into age-specific factors. Higher age groups' openness to AI should be further studied (Cinalioglu et al., 2023). Considering the diverse demographic backgrounds of AI users, it is important to address any potential discrimination and stereotypes during the early stages of development.

5. Conclusions

Providing user-focused prompts and clear instructions improves the quality of AI-generated outputs. This results in content that better aligns with human expectations, enhancing the overall user experience.

Age differences in how people perceive AI content show the need to consider demographics in AI development.

Including human emotions in AI design is crucial for making it more user-friendly and widely accepted across all age groups

References

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. and Amodei, D., Language models are few-shot learners, *Advances in Neural Information Processing Systems*, 33, 1877-1901, 2020.

Cambon, A., Hecht, B., Edelman, B., Ngwe, D., Jaffe, S., Heger, A., Vorvoreanu, M., Peng, S., Hofman, J., Farach, A., Bermejo-Cano, M., Knudsen, E., Bono, J., Sanghavi, H., Spatharioti, S., Rothschild, D., Goldstein, D.G., Kalliamvakou, E., Cihon, P., Demirer, M., Schwarz, M. and Teevan, J., Early LLM-based tools for enterprise information workers likely provide meaningful boosts to productivity, MSFT Technical Report (MSR-TR-2023-43), *Microsoft*, 2023.

Carstensen, L.L., Fung, H.H. and Charles, S.T., Socioemotional selectivity theory and the regulation of emotion in the second half of life, *Motivation and Emotion*, 27, 103-123, 2003.

Cinalioglu, K., Elbaz, S., Sekhon, K., Su, C.L., Rej, S. and Sekhon, H., Exploring differential perceptions of artificial intelligence in health care among younger versus older Canadians: results from the 2021 Canadian Digital Health Survey, *Journal of Medical Internet Research*, 25, e38169, 2023.

Czaja, S.J., Charness, N., Fisk, A.D., Hertzog, C., Nair, S.N., Rogers, W.A. and Sharit, J., Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE), *Psychology and Aging*, 21(2), 333, 2006.

Ekman, P., An argument for basic emotions, *Cognition & Emotion*, 6(3-4), 169-200, 1992.

Epstein, Z., Levine, S., Rand, D.G. and Rahwan, I., Who gets credit for AI-generated art?, *Iscience*, 23(9), 2020.

Gabriel, I., Artificial intelligence, values, and alignment, *Minds and Machines*, 30(3), 411-437, 2020.

Gasper, K., Spencer, L.A. and Hu, D., Does neutral affect exist? How challenging three beliefs about neutral affect can advance affective research, *Frontiers in Psychology*, 10, 2476, 2019.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., Generative adversarial nets, *Advances in Neural Information Processing Systems*, 27, 2014.

Kitayama, S., Mesquita, B. and Karasawa, M., Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States, *Journal of Personality and Social Psychology*, 91(5), 890-903, 2006.

Lee, Y.K., Jung, Y., Lee, I., Park, J.E. and Hahn, S., Building a psychological ground truth dataset with empathy and theory-of-mind during the COVID-19 pandemic, In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43(43), 2582-2587, 2021.

Liu, V. and Chilton, L.B., Design guidelines for prompt engineering text-to-image generative models, In *Proceedings of the 2022*

CHI Conference on Human Factors in Computing Systems (pp. 1-23), 2022.

Markus, H.R. and Kitayama, S., Culture and the self: Implications for cognition, emotion, and motivation, *College Student Development and Academic Life* (pp. 264-293), Routledge, 2014.

Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G. and Sutskever, I., Learning transferable visual models from natural language supervision, *Proceedings of the 38th International Conference on Machine Learning*, *PMLR*, 139, 8748-8763, 2021.

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M. and Sutskever, I., Zero-shot text-to image generation, *Proceedings of the 38th International Conference on Machine Learning, PMLR*, 139, 8821-8831, 2021.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B., High-resolution image synthesis with latent diffusion models, In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10684-10695), 2022.

Roose, K., An AI-Generated Picture Won an Art Prize. Artists Aren't Happy, *The New York Times*, https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html (retrieved March 25, 2024).

Slack, G., What DALL-E Reveals About Human Creativity, *Stanford University Human-Centered Artificial Intelligence*. https://hai.stanford.edu/news/whatdall-e-reveals-about-human-creativity (retrieved March 25, 2024).

Author listings

Yoon Kyung Lee: yoonlee78@snu.ac.kr
Highest degree: PhD in Cognitive Psychology
Position title: Senior Researcher, Department of Psychology, Seoul National University
Areas of interest: Emotion, Natural Language Processing, Affective Modeling, Artificial Intelligence, Human-AI/Robot Interaction, Smart Aging

Yong-Ha Park: peter3061503@snu.ac.kr
Highest degree: M.A. Candidate in Cognitive Psychology
Position title: M.A. Candidate, Department of Psychology, Seoul National University
Areas of interest: Cognitive Psychology, Creativity, Virtual Reality

Sowon Hahn: swhahn@snu.ac.kr
Highest degree: PhD in Cognitive Psychology
Position title: Professor, Department of Psychology, Seoul National University
Areas of interest: Human Cognition, Cognitive Neuroscience, Affective Science, Smart Aging, Human Factors, Human-Robot Interaction