Human Factors in Architecture, Sustainable Urban Planning and Infrastructure, Vol. 191, 2025, 65–74
https://doi.org/10.54941/ahfe1006610

Neural Correlates of Architectural Interior Preferences and Single-Trial Preference Prediction

Yuhong Chen[1,2], Qixu Xie[1,2], and Li Zhang[1,2]

[1]School of Architecture, Tsinghua University, Beijing, 100084, China
[2]Urban Ergonomics Lab, Tsinghua University, Beijing, 100084, China

ABSTRACT

Preference is an important indicator of architectural design quality and human well-being. Current interior space design mainly relies on the designer's subjective judgment and lacks an objective basis. This study aims to quantify event-related potential (ERP) features of architectural interior preference, and examine whether we can infer human preference from single-trial ERP using machine learning. Thirty-six university students participated in an experiment where they viewed architectural interior images and rated them based on their preferences. Significant voltage differences were observed in particular channels (mainly in Oz, O2, Pz, Fp1, Fp2, T7) when participants viewed liked versus disliked images. Source localization indicated that liked images primarily activated the left frontal cortex, while disliked images predominantly activated the left occipital lobe. The within-subject models significantly outperformed the chance level, while the cross-subject models did not show significant results. Also, we found that some visual features can be decoded better than other features by EEG. These findings shed new light on understanding the difference in ERP of interior preference and illustrate the potential for developing a brain-computer interface (BCI) for rapid design evaluation.

Keywords: Architectural preference, EEG (electroencephalography), ERP (event-related potential), Preference prediction, Machine learning

INTRODUCTION

Architectural interior not only serves as functional space but also significantly impacts the psychological and emotional experiences of individuals. Therefore, evaluating human preference for interior is of great importance, as it helps elevate the quality of certain interior designs and satisfies users' needs. Previous research has uncovered brain activity patterns associated with various aesthetic tasks. For example, an fMRI study found that the architectural feature "Fascination" was associated with neural activity in the right lingual gyrus (Coburn et al., 2020). Another study found that architecture with curvilinear features activated the anterior cingulate cortex in beauty judgments (Vartanian et al., 2013). Although significant progress has been made in understanding how the brain processes various stimuli, the specific neural activity patterns associated with architectural interior preference are still not well comprehended.

© 2025. Published by AHFE Open Access. All rights reserved.

Among the neuroimaging techniques, EEG has a high temporal resolution, which allows for capturing real-time changes. Also, EEG is more portable and easier to use in daily scenes compared to other neuroimaging methods. Due to these advantages, EEG is widely used as an objective way to study human mental state in built environment studies (Bower et al., 2019) and preference studies (Jacobsen and Höfel, 2001). In recent years, with the development of machine learning, researchers have begun to train EEG-based classifiers, aiming to unlock the potential of brain-computer interfaces (BCIs) that are of more application value (Akter et al., 2022; Lawhern et al., 2018). Compared to long-term measurement, event-related potential (ERP, often < 2s) is a better way to evaluate various designs quickly. However, it remains to be tested whether ERP can predict human interior preference.

Based on the considerations mentioned above, the main objectives of this study are: 1) To describe the EEG characteristics of the liked and disliked architecture stimuli, offering insight into the neural difference between preferred and non-preferred conditions. 2) To explore the possibility of developing predictive models for architectural interior preference using ERP, exploring the possibility of BCI for rapid architectural design evaluation. Through this research, we aim to offer theoretical foundations and technological tools for evaluating human preference for the built environment.

EXPERIMENT

Materials

The experimental materials for this study consisted of 500 living room interior images generated by DALL-E 3 ("DALL·E 3," n.d.). We chose to use AI-generated images instead of photos or renderers due to the following reasons: 1) Diversity: AI-generated images offer a vast number of novel designs. 2) Unfamiliarity: Since the images are newly generated, they ensure that participants have no prior exposure, eliminating familiarity bias. 3) Controllability: It can ensure that the size and resolution of the image are consistent and the quality is relatively controllable.

We computed six features that may affect architectural preference: (1) average hue, (2) average brightness, (3) average saturation, (4) complexity, (5) sky ratio, and (6) greenery ratio. The complexity was calculated as the ratio of the edge pixels by the Canny edge detection algorithm. The sky ratio and greenery ratio were extracted by the InternImage segmentation model (Wang et al., 2023) pre-trained on the Ade20k dataset (Zhou et al., 2017). We divided the images into binary categories. For average hue, greater than 70 was classified as cool, less than or equal was classified as warm. For the other five features, those greater than the median value were classified as high, and the left were classified as low. Figure 1 displays two examples.

Participants

Thirty-eight university students participated in the experiment (19 males and 19 females). The participants' average age is 22.63. 16 participants major in architecture and other visual arts-related fields. All participants reported no color blindness or weakness, nor history of neurological or psychiatric disorders.

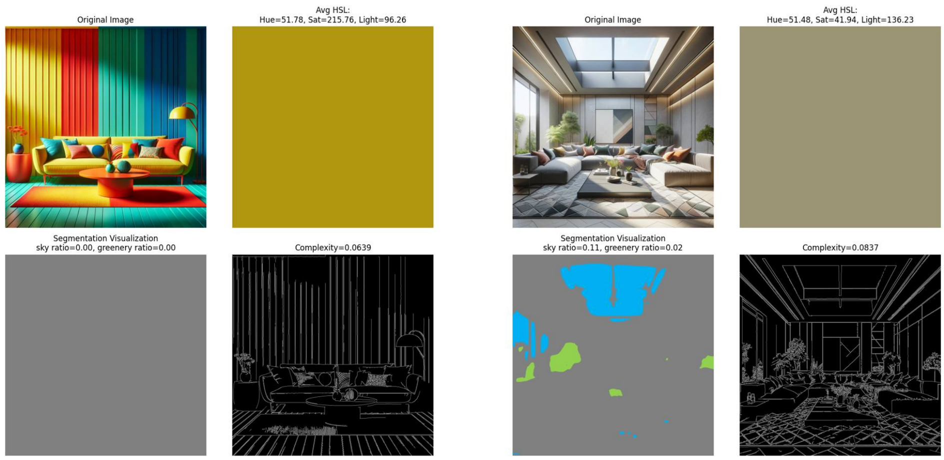

Figure 1: Experiment material examples.

Experimental Procedure

The experiment environment is displayed in Figure 2A. The study used the Neoroelectrics Enobio 32-channel EEG device, which employs AgCl dry electrodes, with a sampling rate of 500 Hz and a bandwidth of 0–125 Hz. The 32 electrodes were placed following the 10–10 system in the places shown in Figure 2B. A reference electrode was placed on the right earlobe. The experiment was conducted in an electromagnetic and sound-shielding room to prevent potential electromagnetic and noise interference. The experimental stimuli were presented on a 24-inch LED screen with a refresh rate of 144 Hz.

First, participants viewed 20 images to familiarize themselves with the experimental procedure. During each trial, the computer interface was set to gray. After a 1500–2000 ms inter-trial interval (ITI), during which a black cross appeared on the screen, the image was presented for 2000 ms, and participants rated the image using the keyboard. Since aesthetic experiences typically diminish with repeated viewings, this study adopted a single-trial design, where each image was presented only once. After familiarizing with the procedure, participants viewed 480 images in random order, with a break of at least 1 minute after every 60 images.

The experiment procedure is shown in Figure 2C. This experiment was approved by the ethics committee (approval number: THU-04-2024-15).

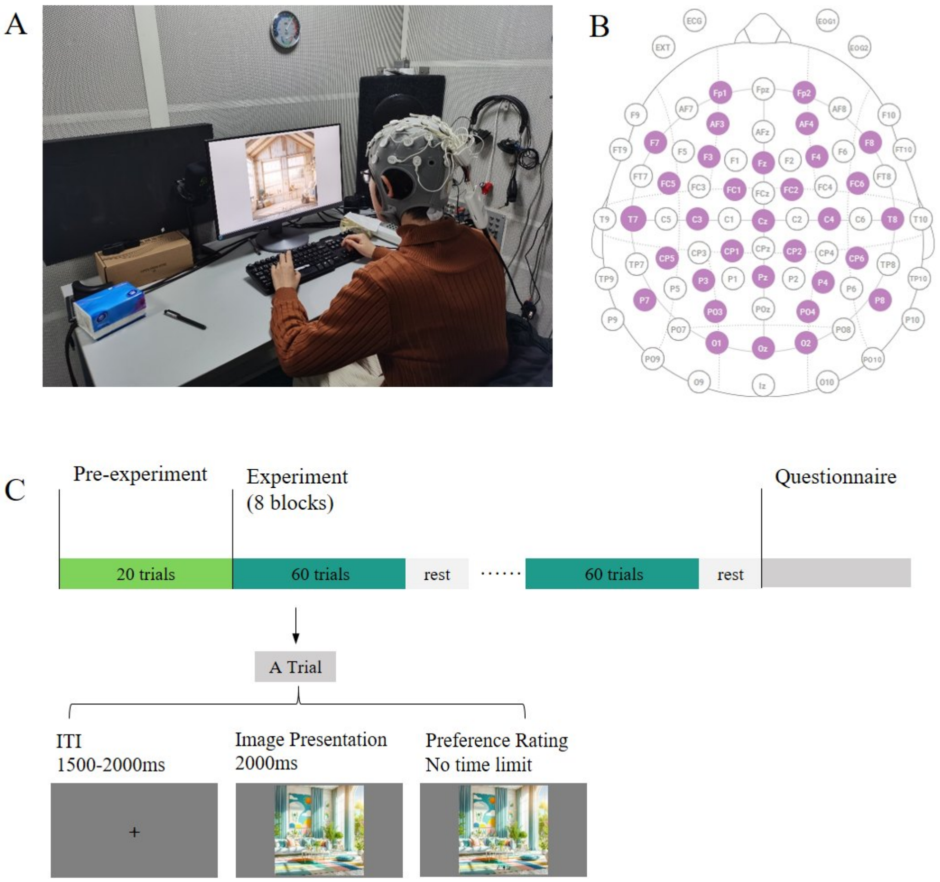

Figure 2: Experiment design. (A) Photo of a participant engaging in the experiment. (B) 32 channels location of the enobio32 EEG device (the channels in purple) (Neuroelectrics, 2022). (C) Experimental procedure.

RESULTS

All the analysis was done using Python 3.11.5 and MNE 1.6.0. We preprocess the data: 50 Hz noise filter, re-referencing to average, interpolating the bad electrodes, independent components analysis (ICA), removing eye movement artifacts, 0.2–10 Hz bandpass filter, baseline correction from –0.5 to 0s, removing trials that exceed 80uV, resampling to 50 Hz. Two participants' data were removed due to too few remaining trials (<100), so the valid sample size is thirty-six.

Temporal Statistical Analysis

We conducted Wilcoxon signed-rank test on the averaged EEG voltage values for the liked and disliked stimuli at each time point and each channel. Figure 3 shows the r value (Figure 3A) and p value (Figure 3B) of the Wilcoxon signed-rank test on the EEG voltage values for liked and disliked stimuli. After 100 ms, many channels exhibit significant differences between the liked and disliked stimuli. In some of the occipital and parietal regions (e.g., Oz, O2, Pz), the liked stimuli exhibit significantly stronger positive voltage than for disliked stimuli. Conversely, in some of the frontal and temporal regions (e.g., Fp1, Fp2, T7) the liked stimuli exhibit significantly stronger negative voltage than for disliked stimuli. We further plotted the waveforms of the channels with more than 30% of the time points having p-values < 0.05 in the 0–1000ms window (Figure 3C).

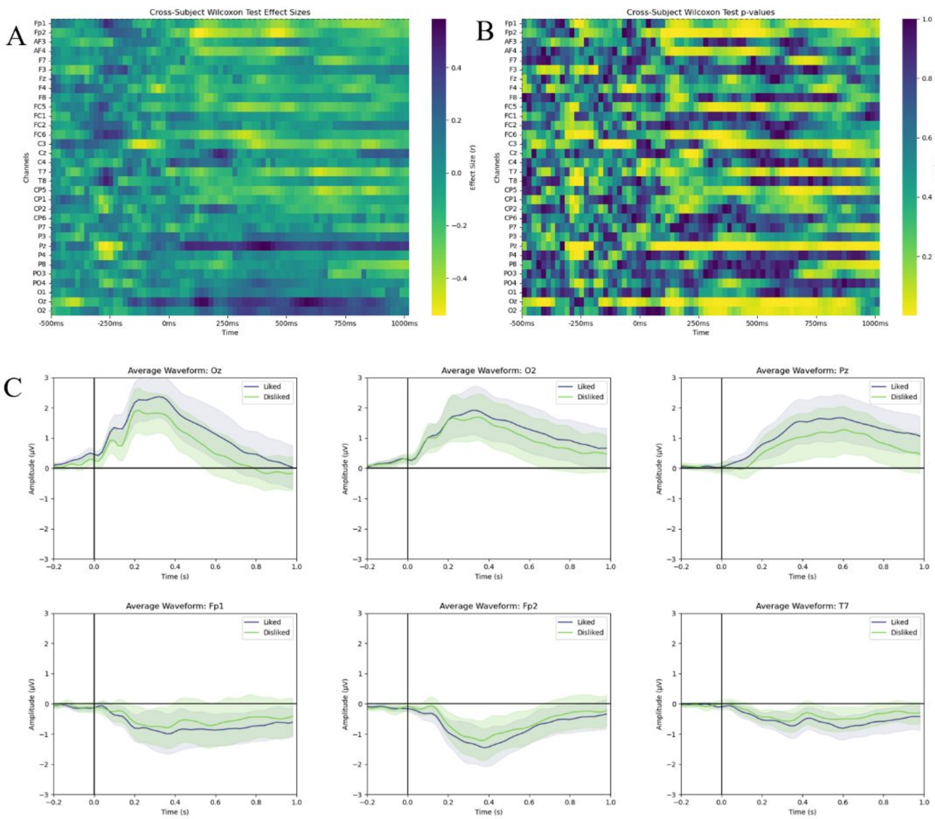

Figure 3: (A) r value and (B) p value of Wilcoxon signed-rank test on the EEG voltage values for liked and disliked stimuli. (C) Waveforms of six channels that more than 30% time points exhibit significant differences in the 0–1000 ms window. The line is the mean of all the subjects, and the shaded areas is the 95% confidence intervals.

Source Localization Analysis

To investigate which brain regions were activated by interior preferences, we conducted source localization analysis. Source localization can infer the origin of EEG signals using inverse problem algorithms, helping to identify the active brain regions. We determine the time window based on the results of the previous Wilcoxon signed-rank test (100–900 ms). We utilized the sLORETA (standardized low-resolution brain electromagnetic tomography) algorithm to determine the activated brain regions for each participant during the liked and disliked trials. The results (Figure 4) showed that in the liked trials, the left frontal cortex was mainly activated, whereas in the disliked trials, the left occipital lobe was mainly activated.

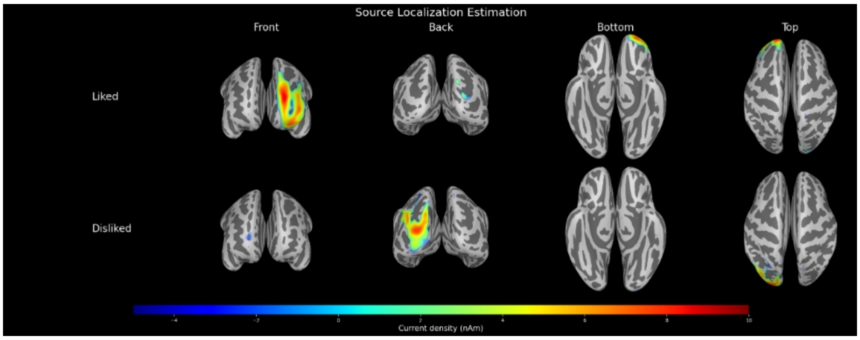

Figure 4: Source localization results.

Machine Learning Decoding Model

We tried to build within-subject and cross-subject machine learning models to see if we could predict the participants' preferences for interior images based on their EEG. For the within-subject decoding, we used the classic logistic regression classifier for binary classification. The input is the voltage value of a trial in all the channels (50 time points * 32 channels, reshaped to one-dimensional vector). Stratified 5-fold cross-validation, was employed to evaluate the model. We also tried to use EEGNet (Lawhern et al., 2018) to train a within-subject classification model. When using EEGNet, we used a 1–40 Hz bandpass filter and downsampled the data to 128Hz during preprocessing, and four-fold cross-validation was used. For model evaluation, we use the area under the receiver operating characteristic curve (hereafter referred to as AUC). For comparison, we shuffled the labels of the data and conducted model training, evaluation and testing in the same manner. Then, we performed paired Wilcoxon signed rank tests to compare the performance of the model on the original data and on the shuffled data.

Figure 5 shows that the within-subject classification model performed significantly better on the original data than the label-shuffled data on the AUC metric (Wilcoxon signed-rank test on the logistic regression results: $W = 145$, $n = 36$, $p = 0.002$, $r = 0.492$; Wilcoxon signed-rank test on the EEGNet results: $W = 143$, $n = 36$, $p = 0.002$, $r = 0.498$). This indicates that EEG signals can reflect preferences to a certain extent. However, we also observed considerable individual differences. Some participants' AUC exceeded 65%, while others had relatively lower performance. We also find that the logistic regression model using only 12 channels' data (Fp1, Fp2, AF4, FC5, FC6, C3, T7, CP5, Pz, P8, Oz, O2) performed significantly better on the original data than the label-shuffled data (Wilcoxon signed-rank test: $W = 159$, $n = 36$, $p = 0.005$, $r = 0.456$). This indicates that 12 channels EEG signals contain information related to interior preferences.

For the cross-subject model, we also tested logistic regression and EEGNet, employing leave-one-subject-out cross-validation. The binary classification model on the original data and the label-shuffled data did not show a significant difference on the AUC metric, indicating that the ERP related to interior preferences of different subjects shows some inconsistency.

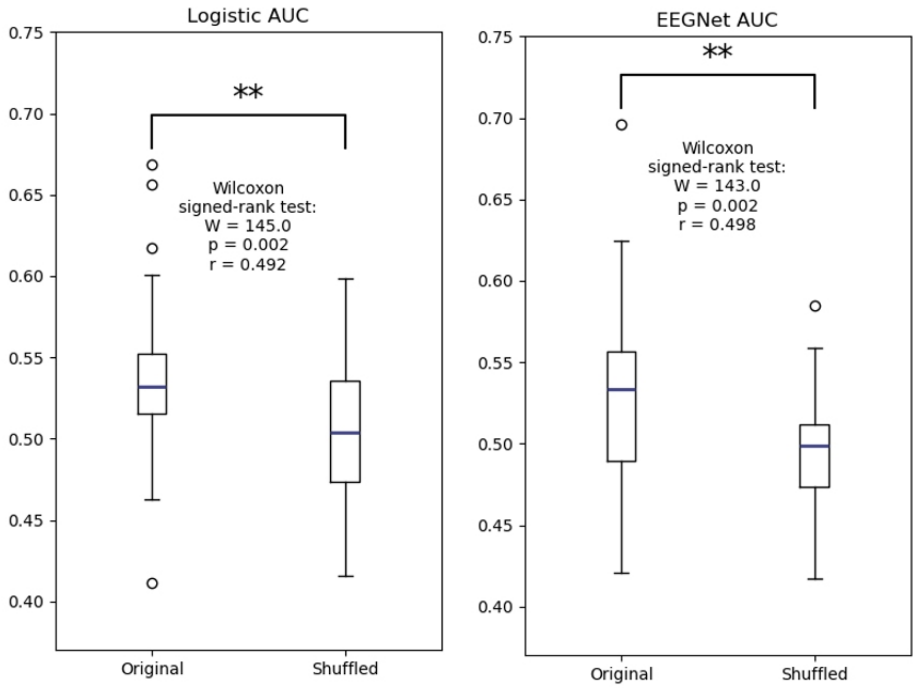

Figure 5: Within-subject classification results using all the channels.

Decoding Architectural Features

Since interior preference may be related to specific architectural features, we also explored the predictive power of EEG for different architectural features. We trained within-subject logistic regression binary classifiers for each feature and compared their performance. The results (Figure 6) indicate that saturation, lightness, complexity and sky ratio can be decoded by EEG signal to some extent, while hue and greenery ratio can not.

Given that architectural features and preferences may be correlated to a certain extent and thus affect EEG, we also conducted Pearson correlation analysis of the six architectural features and preferences. The result (Figure 7) shows that average lightness and sky ratio are significantly positively correlated to preference, while average hue, average saturation and complexity are significantly negatively correlated to preference. Among them, the four features with high significance related to preference for Pearson can also be better classified using EEG (saturation, lightness, complexity and sky ratio, $p < 0.001$). This may indicate that the ability of EEG to decode preference may originate from some basic visual features.

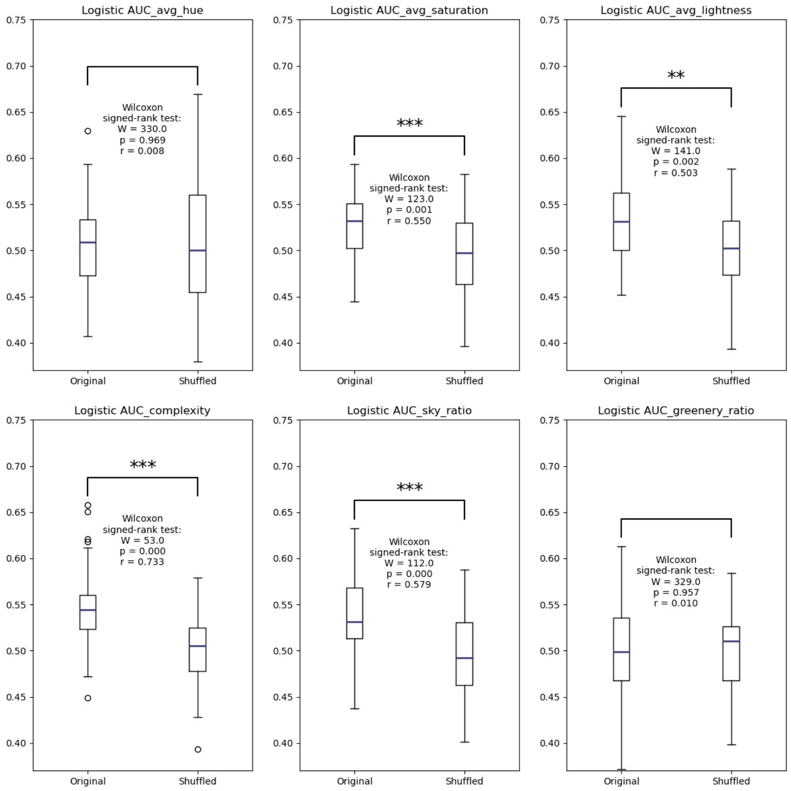

Figure 6: Within subject classification results on 6 features.

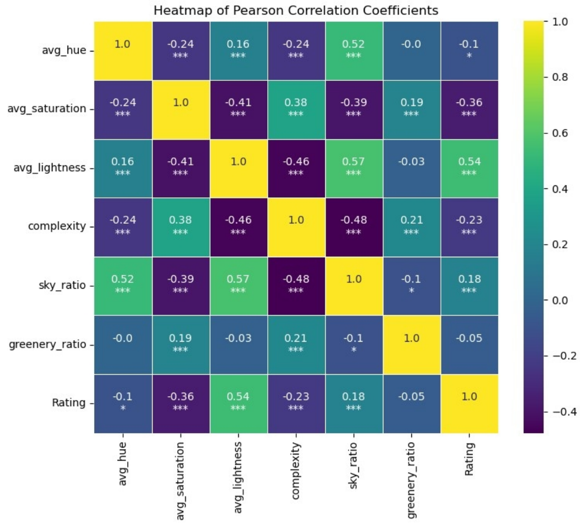

Figure 7: Heatmap of Pearson correlation coefficients between features.

DISCUSSION

This study explored the ERP characteristics elicited by preferred and non-preferred architectural interior images. In general, the findings of this study support the notion that, compared to disliked interiors, liked interiors produce higher levels of visual and cognitive processing within 100–900 ms. We observed strong late positive potential (LPP) in the occipital and parietal areas after 300 ms, and previous studies find that high arousal stimuli elicit LPP (Olofsson et al., 2008). This may suggest that pleasant interior images induce stronger emotional arousal.

Our experiment demonstrates the potential of EEG in estimating human preference toward architectural interiors. The within-subject classification model performed significantly better than the chance level, and some participants' AUC exceeded 65%. It is important to note that the classification performance was highly dependent on the characteristics of the image dataset. In some studies, stimuli with extreme differences are often used in order to elicit greater physiological differences and thus achieve higher classification accuracy. However, a previous review has pointed out that if a small number of extreme stimuli are used, certain features may interfere with the experimental results (Olofsson et al., 2008). Therefore, we believe the diversity of our stimuli can avoid this problem.

This study uses images as stimuli, which provide a controlled environment for quick assessment. Future studies could employ VR or video stimuli to enhance immersion and capture the dynamic aspects of interior spaces more effectively. Also, current non-invasive dry electrode EEG technology may not be very reliable, and the combination of EEG equipment and VR still faces many challenges. Future research should address these problems and improve the applicability of BCI in architecture design.

CONCLUSION

This study provides a foundational understanding of the neural mechanisms underlying architectural interior preferences and examines the possibility of developing a BCI for design evaluation. The key findings are as follows:

1) For the ERP patterns when participants viewed interiors they liked compared to those they disliked, significant differences in voltage were observed in certain channels (mainly Oz, O2, Pz, Fp1, Fp2, T7). The liked images mainly activated the left frontal lobe, indicating higher cognitive engagement.
2) For single-trial decoding of preference, the within-subject binary classification model performed significantly better than random chance, demonstrating the potential of predicting interior preferences from EEG signals.
3) Some visual features (average saturation, average lightness, complexity and sky ratio) can be decoded better than some other features by EEG.

Although the application of EEG in design evaluation is still immature, preferences and some visual features are already partly measurable. This study tested a novel possibility to decode users' preferences through their

EEG signals. This approach allows for the iterative development of designs that more closely align with user preferences. Future research should focus on enhancing the ecological validity of these findings by incorporating more immersive stimuli with advanced neuroimaging tools and addressing individual differences.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of China (No. 52278023) and Cyrus Tang Foundation.

REFERENCES

Akter, S., Prodhan, R. A., Pias, T. S., Eisenberg, D., Fresneda Fernandez, J., 2022. M1M2: Deep-Learning-Based Real-Time Emotion Recognition from Neural Activity. Sensors 22, 8467. https://doi.org/10.3390/s22218467

Bower, I., Tucker, R., Enticott, P. G., 2019. Impact of built environment design on emotion measured via neurophysiological correlates and subjective indicators: A systematic review. Journal of Environmental Psychology 66, 101344. https://doi.org/10.1016/j.jenvp.2019.101344

Coburn, A., Vartanian, O., Kenett, Y. N., Nadal, M., Hartung, F., Hayn-Leichsenring, G., Navarrete, G., González-Mora, J. L., Chatterjee, A., 2020. Psychological and neural responses to architectural interiors. Cortex 126, 217–241. https://doi.org/10.1016/j.cortex.2020.01.009

DALL·E 3 [WWW Document], n.d. URL https://openai.com/index/dall-e-3/ (accessed 7.12.24).

Jacobsen, T., Höfel, L., 2001. Aesthetics Electrified: An Analysis of Descriptive Symmetry and Evaluative Aesthetic Judgment Processes Using Event-Related Brain Potentials. Empirical Studies of the Arts 19, 177–190. https://doi.org/10.2190/P7W1-5F1F-NJK9-X05B

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., Lance, B. J., 2018. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. J. Neural Eng. 15, 056013. https://doi.org/10.1088/1741-2552/aace8c

Neuroelectrics, 2022. Neuroelectrics® Enobio: Instructions for use, Version 4.0. ed. Neuroelectrics Barcelona SLU, Barcelona, Spain.

Olofsson, J. K., Nordin, S., Sequeira, H., Polich, J., 2008. Affective picture processing: An integrative review of ERP findings. Biological Psychology 77, 247–265. https://doi.org/10.1016/j.biopsycho.2007.11.006

Vartanian, O., Navarrete, G., Chatterjee, A., Fich, L. B., Leder, H., Modroño, C., Nadal, M., Rostrup, N., Skov, M., 2013. Impact of contour on aesthetic judgments and approach-avoidance decisions in architecture. Proc. Natl. Acad. Sci. 110, 10446–10453. https://doi.org/10.1073/pnas.1301227110

Wang, W., Dai, J., Chen, Z., Huang, Z., Li, Z., Zhu, X., Hu, X., Lu, T., Lu, L., Li, H., Wang, X., Qiao, Y., 2023. InternImage: Exploring large-scale vision foundation models with deformable convolutions.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., Torralba, A., 2017. Scene parsing through ADE20K dataset, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI, pp. 5122–5130. https://doi.org/10.1109/CVPR.2017.544