From Brainwaves to Understanding: A Study on EEG-Based Communication Systems

Umme Farhana (umme.farhana.23@shizuoka.ac.jp)
Junya Morita (j-morita@inf.shizuoka.ac.jp)
Department of Informatics, Graduate School of Integrated Science and Technology, Shizuoka University,
3-5-1, Johoku, Naka-ku, Hamamatsu, Shizuoka, 432-8011, Japan

Abstract

This research investigates the feasibility of a communication system using Electroencephalography (EEG) signals mediated by a deep learning model, designed to aid individuals with significant communication challenges. The study builds on prior work in EEG-based image generation, testing the accuracy and efficiency of the system in conveying intended meanings. In the experiment, participants were presented images generated from EEG signals and were asked to give titles according to their interpretation. These titles were analyzed using text embeddings derived from a large language model (LLM) to measure cosine similarity. Results indicate that while sender and receiver interpretations often diverged, consistent patterns emerged within and between receivers. This suggests that repeated communication trials will align interpretations over time, improving mutual understanding. The findings highlight the potential of the method to facilitate adaptive communication, though further research is required to optimize its reliability, scalability, and practical applicability.

Keywords: Communication, dream-diffusion, EEG, cosine similarity, Cross-cultural analysis

Introduction

Imagine an individual with a vibrant inner world brimming with ideas, emotions, and creativity but constrained by communication disabilities (such as autism, Down syndrome or just anyone in a new language environment), making it difficult to express these rich internal processes through traditional verbal or written language. For such individuals, this inability to connect with others not only limits their participation in societal activities but also risks alienation, isolation, and loss of their role within the community. Addressing these challenges requires innovative approaches that go beyond conventional communication methods, offering alternative ways to bridge the gap between internal cognition and external expression.

In response to this issue, researchers are developing communication systems that bypass language using Brain-Machine Interfaces (BMIs) and Brain-Computer Interfaces (BCIs). EEG, a non-invasive brain activity measurement tool, shows promise in this field. Advances in AI, especially deep learning, have enabled systems that translate EEG signals into images, externalizing mental processes.

Studies like EEG-based telepathic communication systems (Rao et al., 2014) or on brain-to-image models (Spampinato et al., 2017) represent significant milestones in this area. These systems demonstrate that complex mental representations can be transformed into interpretable outputs, paving the way for novel communication paradigms.

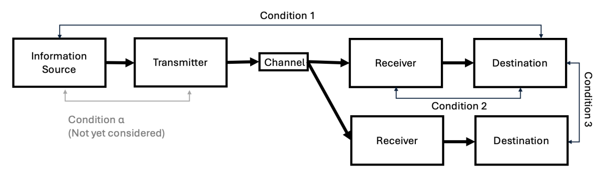

Figure 1: The Shannon-Weaver Model diagram

However, despite this progress, many existing EEG-mediated systems are evaluated using standardized tests that focus on accuracy metrics without considering the dynamic, interactive nature of real-world communication. Communication, whether verbal or non-verbal, is inherently a social process involving the exchange of meaning between individuals. Both cognitive and computer sciences traditionally define communication as a repetitive linear process comprising six elements: a source (who creates the message), a transmitter (which encodes it), a channel (through which the message is sent), a receiver (who decodes it), a destination (who ultimately interprets it), and noise (any interference along the way) (Shannon & Weaver, 1949). In theoretical communication, these elements function linearly, but in real-life situations, they are often more complex and overlapping. Therefore while this framework provides a foundational understanding, effective communication depends on additional qualitative factors, particularly in systems designed for individuals with communication disabilities. To address this gap, we propose evaluating EEG-based communication systems using a set of conditions that align with real-world communicative dynamics (as also shown in Figure 1):

- Conditions 1: There should be similarity between the communicated concept and the interpreted concept.
- Conditions 2: The interpretation of the communicated message should be consistent within the same receiver.
- Conditions 3: The interpretations made by different receivers in a same community should be consistent within each other.
- Condition 4: There should be consistency within the sender while communicating a concept.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

Building on this test framework for novel communication systems, our study applies these conditions to evaluate an existing EEG-mediated communication system. Specifically, we conducted an experiment in which participants were shown images generated from EEG signals and asked to caption these images without any knowledge of the sender's original intentions. By analyzing the captions, we tested hypotheses related to the alignment of interpretations across senders and receivers, as well as consistency within and between participants. The results of this study provide insights into the practical feasibility of EEG-mediated communication systems and their potential to foster meaningful exchanges for individuals with significant communication challenges. By emphasizing real-world communicative conditions—such as the need for mutual understanding, adaptability, feedbacks, and interpretability of outputs (which are often overlooked in standard accuracy-based evaluations), we aim to bridge the gap between technical performance and user-centered outcomes, contributing to the development of systems that truly enhance human connection.

Related Studies

This study builds upon previous research in cognitive science concerning the emergence of communication and the development of EEG-based communication devices. The following sections provide an introduction to both of these foundational research areas.

Emergence of Communication

In the field of cognitive science, communication has been a primary topic of research, with numerous experimental and computational studies exploring its emergence, evolution, and mechanisms. This study follows the tradition of investigating the emergence of communication systems, particularly in contexts where conventional language is absent or limited.

One foundational study (Fay, Walker, Swoboda, & Garrod, 2018) examined the processes by which individuals develop shared symbols for communication. The research highlights the importance of interaction and feedback in shaping shared symbolic systems. Through iterative exchanges between the sender and receiver, shared symbols gradually formed and simplified, demonstrating how new symbolic systems emerge and become conventionalized within groups.

Similarly, another study (Morita, Kojima, Konno, & Hashimoto, 2022) explored the role of individual cognitive traits, particularly autistic traits, in the emergence of novel communication systems. By employing simplified coordination games and the Autism-Spectrum Quotient (AQ), the study found that individuals with autistic traits were more likely to create novel, structured communication systems. This challenges traditional views that associate autistic traits primarily with communication deficits and instead suggests a unique role in language evolution.

Previously, one study (Kirby, Cornish, & Smith, 2008) investigated how language structures evolve through cultural transmission. Their study utilized artificial language learning paradigms to demonstrate how linguistic regularities emerge over generations of learners. The research suggests that iterative learning processes contribute to the stabilization of communication systems, emphasizing the interplay between cognitive constraints and social interaction in shaping communication.

Another relevant study (Galantucci, 2005) explored the formation of graphical communication systems in experimental settings. Participants who lacked a predefined communication system successfully developed novel graphical symbols to convey messages. The study demonstrated that even in the absence of spoken language, humans naturally develop alternative symbolic systems, reinforcing the adaptability of communication.

Furthermore, Roberts and Levinson (2017) examined how interactive alignment and social coordination contribute to the emergence of linguistic structures. Their study found that repeated interactions led to the formation of systematic patterns in communication, with individuals unconsciously adapting their symbolic representations to enhance mutual understanding. This research underscores the role of communicative interaction in driving language evolution and highlights mechanisms relevant to EEG-mediated communication systems.

Together, these studies provide a foundational framework for understanding the emergence of communication systems. They emphasize the dynamic interplay between cognitive traits, interactive feedback, and social learning in shaping symbolic communication. Based on these fundamental studies, the current research aims to validate EEG-mediated communication. We don't assess the replication accuracy of images generated from EEG and generative models. Rather, we assess how much the generated images transmit the meaning of the original concept by viewing them as novel communication systems.

EEG as a communication media

There has been a long-standing tradition of utilizing EEG as a communication medium. One prominent line of research is the development of the P300 speller (Farwell & Donchin, 1988), a brain-computer interface (BCI) technology that enables communication by detecting the P300 wave, an event-related potential in EEG signals. This system has proven particularly beneficial for individuals with severe physical disabilities, such as Amyotrophic Lateral Sclerosis (ALS), by allowing them to select characters from a flashing grid. However, the P300 speller has inherent limitations, including slow communication speed, user fatigue due to prolonged concentration, variability in signal quality, and the requirement for extensive calibration.

We posit that these limitations stem from the reliance on conventional language as a communication medium. Unlike expressing ideas through typed text, studies on the emergence of communication suggest the possibility of developing novel communication systems that bypass linguistic constraints. Recent advances in EEG-based communication sys-

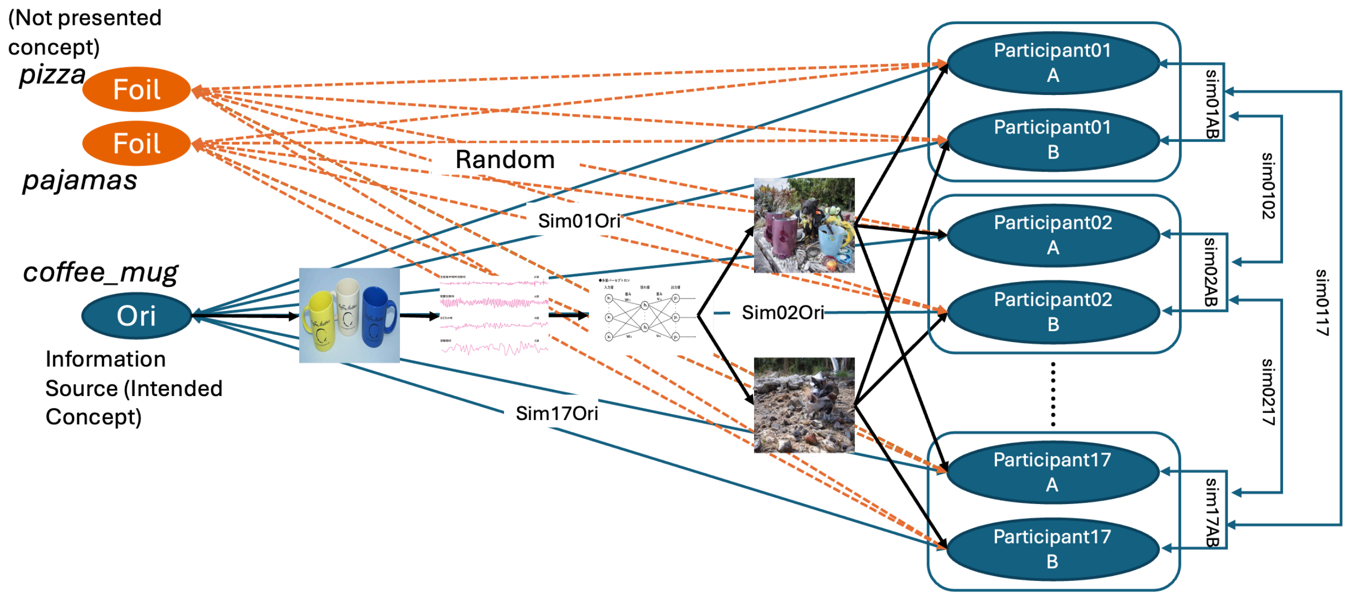

Figure 2: The process of the proposed communication model and the experimental design

tems highlight the potential of generative AI technologies in facilitating such alternative communication methods. One foundational study demonstrated that EEG contains patterns related to visual content, which can be effectively used to generate images that are semantically coherent with the original visual stimuli (Kavasidis, Palazzo, Spampinato, Giordano, & Shah, 2017). This finding, when combined with the theoretical framework proposed by Fay et al. (2018) suggests that even if the generated images are not perfect representations, a shared symbol system can emerge when senders and receivers develop consistent associations between the generated outputs and their interpretations.

The technology for generating images from EEG has advanced significantly, particularly with the development of state-of-the-art (SOTA) models such as DreamDiffusion (Bai et al., 2023). This model leverages pre-trained text-to-image architectures and employs temporal masked signal modeling to pre-train an EEG encoder for producing robust EEG representations. Additionally, DreamDiffusion incorporates the CLIP image encoder (Radford et al., 2021) for further supervision, aligning EEG, text, and image embeddings even with a limited number of EEG-image pairs. However, despite these advancements, existing EEG-based image generation models have not been systematically tested within a communication framework.

Beyond DreamDiffusion, recent research continues to explore alternative EEG-based BCI approaches. For instance, studies have investigated multimodal methods that integrate deep learning architectures to enhance EEG signal interpretation and improve the robustness of BCI-driven image synthesis (Ferrante, Boccato, Bargione, & Toschi, 2024).

Another recent study has proposed an adaptive learning framework to optimize EEG decoding in real time, improving the accuracy of generated representations and reducing individual variability (Guo, 2024). These advancements further reinforce the potential of EEG-driven generative AI in facilitating non-verbal communication.

Given these developments, our study aims to examine EEG-based image generation in a communicative setting, evaluating its feasibility as a novel medium for interaction beyond conventional linguistic expression. The experimental conditions are detailed in the first section and evolve into the following hypotheses.

Hypotheses

We evaluated the capabilities of cutting-edge EEG-based image-generation models as a medium for concept transmission. In conjunction with this exploration, we devised an experimental framework as depicted in Figure 2. At the heart of this framework lies the information transmission sequence for EEG-based image generation, modeled after the Shannon-Weaver framework.

In this sequence, the initial concept (e.g., coffee_mug) is transformed into an image—a process that is conceptualized as mental but is presented to participants in the experiment. Following this conceptualization, an EEG signal is recorded and input into a pre-trained neural network model, which produces a series of images based on different random noise patterns. These generated images are then presented to several receivers, who interpret them independently. Based on this information transmission model and the conditions outlined in section 1, we established the following hypotheses:

- Hypothesis 1: The similarity between the intended concept and the interpreted concept is significantly higher than the similarity between randomly paired concepts (Foiles in the figure). In equation form:

$$\sum_{t=1}^{n}\sum_{i=1}^{m}\frac{Sim(Concept_{t,i},Ori_t)}{n\cdot m} > \text{Random}$$

where t, i, n, and m denote a target concept, an individual interpreting the concept, the number of concepts, and the number of participants, respectively.

- Hypothesis 2: The interpretation of the communicated message is consistent within the same receiver. In equation form:

$$\sum_{t=1}^{n}\sum_{i=1}^{m}\frac{Sim(Concept_{tA,i},Concept_{tB,i})}{n\cdot m} > \text{Random}$$

where tA and tB are interpretations generated from the different images but the same original concepts.

- Hypothesis 3: Recipients with a common cultural or biological background will show a higher degree of similarity in their interpretations of the communicated message. In equation form:

$$\sum_{t=1}^{n}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\frac{Sim(Concept_{t,i},Concept_{t,j})}{\frac{m(m-1)}{2}\cdot n}$$

$$> \sum_{t=1}^{n}\sum_{i=1}^{m}\frac{Sim(Concept_{t,i},Ori_{t,i})}{m\cdot n}$$

where j represents another individual with a shared background with an individual i.

Method

To test the above hypotheses, we gathered interpretations for images generated by an EEG-based image-generation model. The specific details of the experiment are presented as follows.

Materials

A questionnaire was created using Google Forms for this study. It included 20 images presented in Figure 3, which were generated by the DreamDiffusion model (Bai et al., 2023). From the materials presented in this paper, we randomly selected 10 original images and generated two images for the image by contextualizing corresponding EEG signals in the repository[1] and using different random initial seeds. The experiment reported in Bai et al. (2023), the participants generated EEG while presenting the images. The original images were assigned with concepts (original caption) expressed in natural language, which were collected from WordNet (Deng et al., 2009) by matching with the ImageNet Code of the original images (Table 1). Note that the image code "n03063599" (coffee mug in original caption) was duplicated in two images. In the later analysis, we treated these duplicated captions as separate concepts.

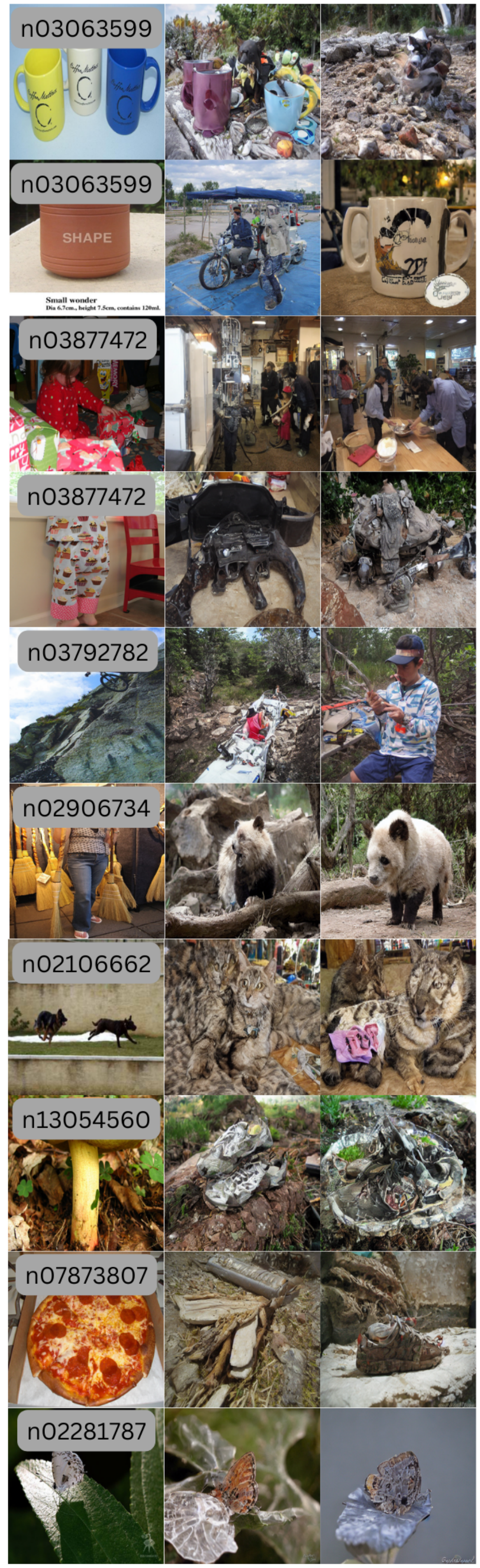

Figure 3: Images used for data collection

[1]https://github.com/bbaaii/DreamDiffusion

Table 1: List of images from ImageNet

Image Code	Original Caption
n03063599	coffee_mug
n03877472	pajama
n03792782	mountain_bike
n02906734	broom
n02106662	German_shepherd
n13054560	bolete
n07873807	pizza
n02281787	lycaenid

Participants and Procedure

Seventeen participants were selected for the study. The group included eight Japanese participants (seven male, one female) and nine Bangladeshi participants (six male, three female). Participants from two nationalities were included to examine Hypothesis 3. All Bangladeshi participants had migrated from Bangladesh to Japan at least one year prior to the study. All participants had a science-related educational background, with the minimum level of education being undergraduate student status. English is second language for all the participants. Two of the participant's responses were recorded in Japanese which were translated into English using Google Translator prior to the analysis. The age range of the participants was 19-38 years with average age of 25.9 years and standard deviation of 4.3 years.

Analysis

The label written by the participants were converted into numerical vectors using OpenAI's embedding API (OpenAI, 2021), allowing for the calculation of cosine similarity between responses.

Result

This section presents the results obtained from the similarity calculations performed in our study.

Hypothesis 1: Sender-Receiver Understanding

To examine hypothesis 1, cosine similarities were computed between the original labels and the responses for the original concept from all 17 participants. Additionally, a dataset of random responses was generated for comparison by shuffling obtained responses within the participants. For example, in the random dataset, the response for "pajama" is connected to "coffee mug" and cosine similarities are calculated cosine similarities between such unmatched pairs.

The results, presented in Figure 4, compare the random similarity with the three participant groups ($Sim(Concept_i, Ori)$ including all participants, $Sim(Concept_Ja, Ori)$ including only Japanese participants and $Sim(Concept_Ba, Ori)$ including only Bangladesh participants). The average values were calculated across all the concepts (EEG-generated images) and all the participants in the group. The attached error bars of Figure 4 were the standard error of the means across concepts (EEG-generated images, $n = 20$). To examine the differences between these similarity values, one-way within analysis of variance (ANOVA) was conducted using each concept (generated image) as a unit of analysis.

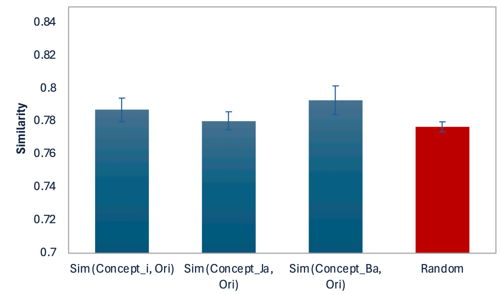

Figure 4: Comparison with the similarity with random response

The result revealed a significant main effect of the similarity values ($F(3,57) = 3.66$, $p < .05$). Multiple-comparison by Holm method reveals the difference between the random and $Sim(Concept_Ba, Ori)$ ($p < .05$). The result indicated that EEG-generated concepts could transmit the original concepts to some extent, at least for a specific participant's group. However, the effect was not so strong, and it remained at a random level for a specific group of participants.

Hypothesis 2: Similarity of Same Pairs

To test the second hypothesis, we analyzed the similarity of captions generated by the same participants for images derived from the same original concept. The within-concept similarity values were calculated between two responses obtained from the EEG-generated images produced by the same original image. The random similarity values were also calculated from pairs of concepts randomly chosen from the same participant. As shown in Figure 5, within-concept similarity scores are consistently higher than random similarity scores across all participants. We also calculated the average across all the participants ($n = 20$) and all the concepts (original images $n = 10$). The obtained scores are 0.824 for the within-concept similarity and 0.776 for the random similarity value, which indicates significant differences ($t(9) = 3.04$, $p < .01$) when we treat the concepts as units of analysis.

As can be seen from the above analysis, the second hypothesis was supported, showing that participants maintain a consistent interpretation when describing images generated from the same concept, reinforcing the stability of conceptual understanding in the captioning for EEG-generated images.

Hypothesis 3: Inter-Participant Similarity

To validate the third hypothesis (response similarity within the group), we analyzed similarity scores between the pairs of responses from the same group. For each response, the similarity with the response that the other participants made the same concept (EEG-generated image) was calculated and

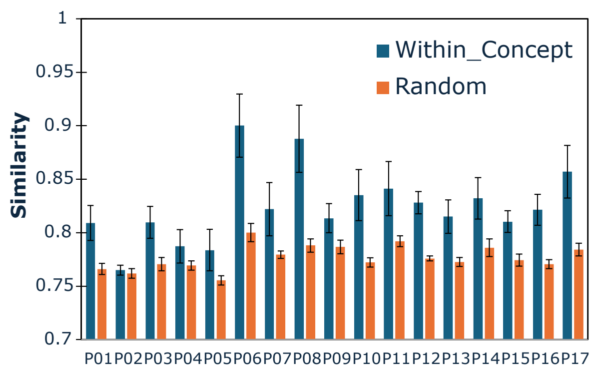

Figure 5: Similarity within the concept

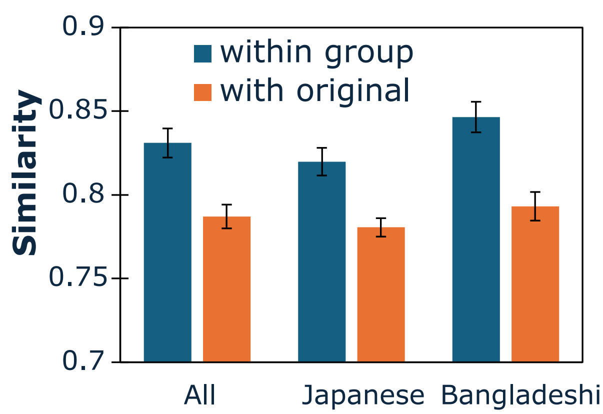

Figure 6: Similarity between the groups

compared with the similarity scores obtained in the first analysis ($Sim(Concept_i, Ori)$). The results are summarized in Figure 6.

Using the values in the figure, we conducted 2 (Pairs: within group vs. with original) x 3 (Groups: all, Japanese, Bangladesh) within ANOVA was conducted. As a result, the significant main effect of the pairs ($F(1,19) = 22.24$, $p < .01$) and the significant main effect of the group ($F(2,38) = 32.37$, $p < .01$) and the interaction between the pairs and the groups ($F(2,38) = 11.297$, $p < .01$) were obtained. The subsequent analyses of the main effects confirmed the simple main effects of pairs in every group. These results suggest that participants align more closely within their cultural groups than with the original labels, emphasizing demographic influences on similarity patterns.

Discussion and Conclusion

This study explores the feasibility of EEG-mediated communication systems, revealing both their potential and current limitations. While EEG-based image generation conveys some conceptual information, sender-receiver interpretations often diverge. However, consistent patterns within and between receivers suggest that iterative refinement could improve alignment.

Statistical analysis supports the hypothesis that intended concepts are more accurately interpreted than random chance in a specific group (hypothesis 1). Additionally, individual receivers maintain stable conceptual mappings (hypothesis 2), and cultural similarities (hypothesis 3) enhance interpretation consistency. These findings highlight the role of cognitive and cultural schema in shaping communication.

Despite these promising results, limitations remain. The first limitation is observation in Japanese participants who did not indicate significant similarity between their response and the original concept, compared to the similarity between random combinations. This reason may be attributed to their language skills. Compared to the international students (Bangladesh) living in their foreign country, Japanese students rarely use English expressions. Such a language related issues may affect the similarity scores calculated by LLM.

The small sample size ($n = 17$) may also limit generalizability, and reliance on text embeddings for similarity measurement introduces potential biases. Future work should refine evaluation metrics, expand participant pools, and incorporate iterative feedback to enhance communication accuracy.

This research contributes to EEG-based communication by systematically testing its viability within a sender-receiver framework. Future advancements should focus on optimizing model reliability, reducing ambiguity, and integrating adaptive learning mechanisms to personalize outputs. Longitudinal studies can further explore how alignment evolves over time.

While the study presents promising preliminary findings, further improvements in machine learning (the accuracy and robustness of EEG-to-image translation models) and cognitive modeling (the interpretability of the generated outputs) are needed to make EEG-mediated communication practical. Bridging the gap between feasibility and real-world application can pave the way for inclusive communication technologies for individuals with impairments.

References

Bai, Y., Wang, X., Cao, Y., Ge, Y., Yuan, C., & Shan, Y. (2023). Dreamdiffusion: Generating high-quality images from brain eeg signals. *arXiv preprint arXiv:2306.16934*. Retrieved from `https://arxiv.org/abs/2306.16934`

Deng, J., Dong, W., Socher, R., Li, L., Li, K., & Fei-Fei, L. (2009). *Imagenet.* (Available: `http://www.image-net.org/`)

Farwell, L., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography & Clinical Neurophysiology*, *70*(6), 510–523.

Fay, N., Walker, B., Swoboda, N., & Garrod, S. (2018). How to create shared symbols. *Cognitive Science*, *42*, 241–269. doi: 10.1111/cogs.12600

Ferrante, M., Boccato, T., Bargione, S., & Toschi, N. (2024). Decoding visual brain representations from electroencephalography through knowledge distillation and latent diffusion models. *Computers in Biology and Medicine*, *178*, 108701. doi: 10.1016/j.compbiomed.2024.108701

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767. doi: 10.1207/s15516709cog0000_34

Guo, H. (2024). EEGVision: Reconstructing vision from human brain signals. *Applied Mathematics and Nonlinear Sciences*, *9*(1), 1–16. doi: 10.2478/amns-2024-1856

Kavasidis, I., Palazzo, S., Spampinato, C., Giordano, D., & Shah, M. (2017). Brain2image: Converting brain signals into images. In *Proceedings of the 2017 acm on multimedia conference (mm '17)* (pp. 1800–1808). doi: 10.1145/3123266.3127907

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686. doi: 10.1073/pnas.0707835105

Morita, J., Kojima, A., Konno, T., & Hashimoto, T. (2022). Influence of autistic traits on the formation of novel symbol communication systems. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, *29*(4), 557–574. doi: 10.11225/cs.2022.024

OpenAI. (2021). *Openai embeddings.* (Available: `https://platform.openai.com/docs/guides/embeddings/`)

Radford, A., Kim, J., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., … Sutskever, I. (2021). Learning transferable visual models from natural language supervision. *arXiv preprint arXiv:2103.00020*. Retrieved from `https://arxiv.org/abs/2103.00020`

Rao, R., Stocco, A., Bryan, M., Sarma, D., Youngquist, T., Wu, J., & Prat, C. (2014). A direct brain-to-brain interface in humans. *PLOS ONE*, *9*(11), e111332.

Roberts, S., & Levinson, S. (2017). Conversation, cognition and cultural evolution: A model of the cultural evolution of word order through pressures imposed from conversational interaction. *Interaction Studies*, *18*(3), 494–517. doi: 10.1075/is.18.3.07rob

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Spampinato, C., Palazzo, S., Kavasidis, I., Giordano, D., Shah, M., & Sorbello, R. (2017). Deep learning human mind for automated visual classification. In *Proceedings of the ieee conference on computer vision and pattern recognition (cvpr)* (pp. 6809–6817).