# EEG VARIATIONS AS A PROXY OF THE QUALITY OF THE DESIGN OUTCOME

**Li, Shumin;**
**Becattini, Niccolò;**
**Cascini, Gaetano**

Politecnico di Milano

## ABSTRACT

This paper presents an EEG (Electroencephalography) study that explores the correlation between the EEG variation across design stages and the quality of the design outcomes. The brain activations of 33 volunteers with engineering backgrounds were recorded while performing a design task using a morphological table to develop an amphibious bike. The EEG variations from the analysing/selecting stage to the illustrating stage were analysed based on the EEG frequency band and channel sets. A significant correlation between the detail level of the design outcome and the power variation mode was observed in theta, alpha and gamma bands, each involving different channel sets. Compared to the assessment results from two evaluators, using EEG variations as a proxy of the detail level of the design outcome could reach a maximum accuracy of 0.727, precision of 0.765, and recall of 0.889. These results also provide suggestions on the selection of the frequency bands and channel sets to achieve better prediction performance according to each metric.

**Keywords**: Electroencephalography, Human behaviour in design, Design cognition, Design Quality, Research methodologies and methods

**Contact**:
Li, Shumin
Politecnico di Milano
Italy
shumin.li@polimi.it

**Cite this article:** Li, S., Becattini, N., Cascini, G. (2023) 'EEG Variations as a Proxy of the Quality of the Design Outcome', in *Proceedings of the International Conference on Engineering Design (ICED23),* Bordeaux, France, 24-28 July 2023. DOI:10.1017/pds.2023.154

# 1 INTRODUCTION

The estimation of the goodness of an ideation or conceptual design process or its outcomes typically is an a posteriori process. Once the generative session is over, evaluators apply specific metrics to determine the outcome quality for, e.g. level of detail, feasibility, or quantity of ideas/fluency. This, indeed, cannot take place during the design process as it would be intrusive to the designers' activities or, even if carried out through non-intrusive observations, this would require the involvement of multiple evaluators to ensure the reliability of the different subjective estimations (again a-posteriori). Then, it is difficult to provide real-time feedback during the design process to improve the quality of the results and the overall process is time-consuming for both the designers and the evaluators (Adolphy et al., 2009).

However, deepening the understanding of cognition is primary to finding possible signatures, precursors or determinants during the design process, such as stimuli, of effective thinking (moves) to shorten the design process and improve the quality of its outcomes. Within design cognition studies, neuroimaging is an emerging technology trend to collect valuable objective data during the design process. Among the available neuroimaging signal acquisition techniques, electroencephalography (EEG) shows promising features as it collects data about brain activity with a high temporal resolution, especially compared to observation-based data collection. EEG-data-based exploration of cognition, with an impact on design, has advanced from simple divergent thinking tests to more complex design tasks (Fink and Neubauer, 2006; Hu et al., 2022; Lukacevic et al., 2022). Plenty of protocol analyses have already reported the potential of using such technology to objectively describe the behaviour of the designer (Jia and Zeng, 2021; Vieira et al., 2022), while the link between the measurements and the final design outcome quality is yet to be discussed.

EEG produces wave-like data whose characteristics emerge in the frequency, in the time domain or both. Several fundamental features, such as the amplitude range and the distinction of frequency bands which normally occurs in human brain activations, have been defined by clinical studies (Stern, 2013). It is also well studied during motor behaviour and in cognitive tasks, the event/task-related change in response to a variety of different stimulus or task-related factors (Pfurtscheller et al., 1993). And it is already been revealed the correspondence between the changing mode to the function depends on the frequency components and brain area (Lopes da Silva, 2006). Experiments on divergent thinking tests correlating creativity or originality to EEG rhythm variations are aligned with these findings (Benedek et al., 2011). Design tasks, however, include both divergent thinking (e.g. concept exploration) and convergent thinking (e.g. concept selection and integration). Whether and how the design outcome quality is correlated to the EEG variation mode across different design stages is not yet studied.

This paper describes a study from a larger research project whose goal is to understand how EEG data can inform the analysis of the design cognition process (Li et al., 2021). This used an experimental protocol that included different tasks entailing different thinking styles and activities to explore the cognitive behaviour from basic creativity tests to the real design task. The study reported in this paper specifically focuses on a design task which uses the morphological table as a tool that requires concept exploration, selection and integration. The assessment of the quality of the design outcome is based on the clarity/level of detail of the related sketches and annotations to address these research questions:

- Can we use EEG (de-)synchronisation (EEG power variation) to predict the clarity/level of detail of the design outcome?
- How many and what are the essential channels, and in which frequency band could we enable the observation of the correlation between EEG variation and the quality of the design outcome?
- How much could we rely on the EEG (de-)synchronisation model to indicate the quality of the design outcome?

To answer these questions, this paper first briefly introduces the background of the field by presenting how the criteria adopted in the current paper got employed by previous studies and their achievements. Then, it illustrates the experimental protocol of the design task, the data acquisition and processing pipeline together with the criteria to statistically verify results. Afterwards, it shows and discusses the preliminary results. The conclusion presents a summary and insights for future development.



https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

## 2 RELEVANT BACKGROUND

### 2.1 The EEG variation

The EEG rhythm, which describes brain neurons' voltage fluctuations within specific frequency bands (Delta [0-4 Hz], Theta [4-7 Hz], Alpha [7-13 Hz], Beta [13-30 Hz] and Gamma [above 30 Hz], has already shown to correlate basic brain functions and with the localisation of brain areas. (Stern, 2013)

The power changes at each rhythm were quantified via task-related power (TRP) by comparing the activation between reference and activation phases (Pfurtscheller and Lopes Da Silva, 1999). Depending on the time length of the phenomenon to observe, different estimators should be used to describe the power variation. For the analysis focusing on changes related to a task within a steady-state process in a time range of 100 ms to seconds, as reported by Hummel and Gerloff (2006), the task-related power increase/decrease (TRPI/TRPD, or task-related (de) synchronisation) is more suitable. On the contrary, studies addressing shorter time frames should investigate event-related (de) synchronisation (ERS/ERD).

### 2.2 EEG-based performance estimation

As aforementioned, different researchers have reported EEG variation associated with performing divergent thinking or several different types of design activities, generally assessing the outcomes using creativity metrics. A period of relaxation by gazing at a fixed point commonly served as the baseline for reference. An EEG TRPI from the baseline, especially in the alpha band, is observed as a proxy of brain activation in most experiments mentioned in previous literature contributions. However, concerning the EEG variation through different stages of the design process, the variation is expected to be more dynamic. Also, the inhibitory state can play an active (inhibitory) role in gating the transfer of information in specific neuronal pathways (Lopes da Silva, 1991).

As creativity is a crucial skill of a designer, its understanding and empowerment via EEG metrics already gathered interest in the field (Jaarsveld et al., 2015; Liu et al., 2018) despite it being still at a preliminary stage. EEG variation, especially the model based on ERS/ERD, already provides different possibilities in clinical diagnosis (Pfurtscheller and Silva, 2017), decision-making prediction (Ratcliff et al., 2009), emotion detection (Lan et al., 2016) and other human-machine interface applications (Foldes and Taylor, 2013).

The performance of any prediction model depends on the accuracy of the prediction. However, imbalanced datasets require other metrics to be considered for such prediction potential and the confusion matrix might be helpful to provide a more comprehensive analysis (Kulkarni et al., 2020). As such, the potential of using EEG variation to predict the quality of the design outcome for achieving different predictive metrics' performances, can be more properly estimated.

### 2.3 The goal of the current research work

Overall, the study aims at investigating the links between EEG variations and the quality of the design outcome, thus highlighting the characteristics of EEG variation usable as a proxy of a design outcome quality. EEG variations are observed during the stage of analysing the contents of a morphological table and selecting partial solutions to combine and the stage of illustrating the final design concept through sketches and annotations. The association between frequency bands and channels to discriminate the quality of the outcome based on the EEG variation mode and to build a predictive model is statistically checked. The confusion matrix for the predictive model describes its performance as accuracy, precision and recall.

## 3 METHOD

### 3.1 Experimental protocol

This contribution exploits EEG data acquired from a larger research protocol, consisting of five different tasks to enable multi-dimensional analysis across different perspectives of design activities. As the first task, all the participants faced a warm-up task. The order of the remaining four ones, including the Design with Morphological Table task (DwMT), is random to avoid fatigue biases.

At the beginning of the experiment, two relaxation sessions are scheduled and recorded: each subject is instructed to first relax and maintain, for 30 seconds, an open-eyed gaze on the central cross that is displayed on the screen. Then, they are asked to relax with their eyes closed for additional 30 seconds. The initial period of observation with open eyes is exclusively employed as the baseline for further



https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

data processing. Both sessions are necessary to validate data acquisition, as the transition from open to closed eyes, without any other cognitive activity, should be coherent with Stern (2013).
To make each participant capable of carrying out the task in an expected manner, the on-screen example of designing a vegetable collection system (Figure 1, left) complemented the instructor's further clarification previously provided. The final solution of the example is not given to avoid potential design fixation (Jansson and Smith, 1991). Upon confirmation of the participant's comprehension of the design process, one could proceed with additional task instructions. These instructions specify that audible notifications will beep twice: at halftime and when it remains 1 minute (Figure 1, centre). As soon as the task begins, the participant can check for the remaining time on screen (Figure 1, right). Participants are provided with a printed A4-pager design brief for developing an amphibious bike (Figure 2). To shift from the examples to the instructions and to the task itself, the subject should press the space button on the keyboard to proceed. The task naturally consists of chronological stages of reading the requirements (stage 1), analysing and selecting the partial solutions (stage 2), and then sketching (stage 3). No time limitation constrains the execution of the task.
All the experiment sessions are audio/video recorded to facilitate body movement detection that triggers EEG artefacts and stage recognition. Since the whole experimental process is programmed with Psychopy 3 (Peirce et al., 2022), the application's log file enables the synchronisation of the different datasets.

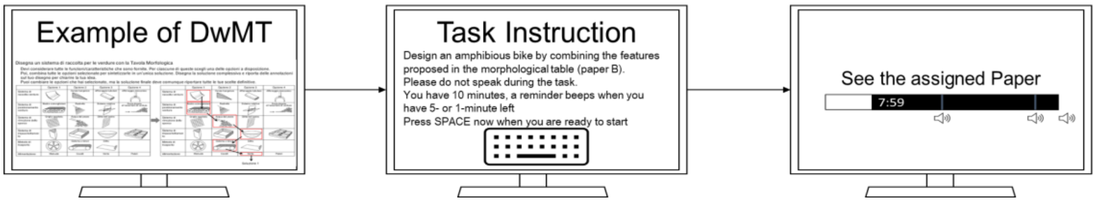
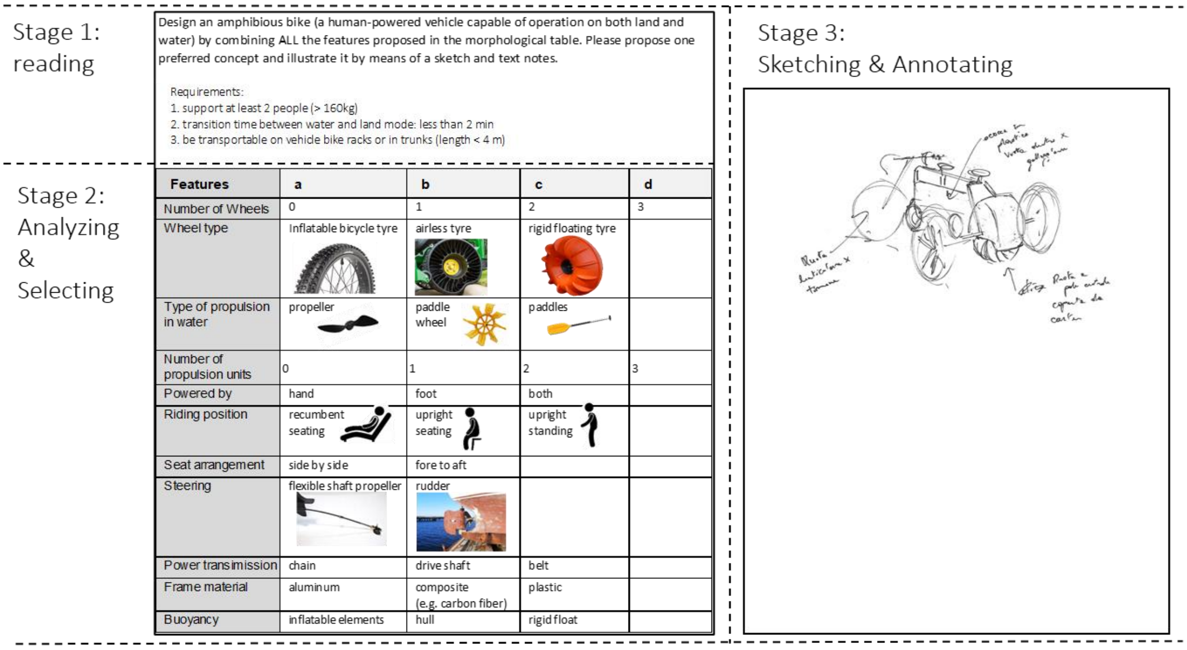



*Figure 1. Instruction of Design with Morphological table on screen*




| Features | a | b | c | d |
|---|---|---|---|---|
| Number of Wheels | 0 | 1 | 2 | 3 |
| Wheel type | Inflatable bicycle tyre | airless tyre | rigid floating tyre | |
| Type of propulsion in water | propeller | paddle wheel | paddles | |
| Number of propulsion units | 0 | 1 | 2 | 3 |
| Powered by | hand | foot | both | |
| Riding position | recumbent seating | upright seating | upright standing | |
| Seat arrangement | side by side | fore to aft | | |
| Steering | flexible shaft propeller | rudder | | |
| Power transmission | chain | drive shaft | belt | |
| Frame material | aluminum | composite (e.g. carbon fiber) | plastic | |
| Buoyancy | inflatable elements | hull | rigid float | |

*Figure 2. The morphological table for designing an amphibious bike & design stage division*

## 3.2 Criteria for the assessment of design outcomes and clustering

The DwMT task returns a sketch with annotations of the final design of the amphibious bike. Two independent evaluators rated them with 4-level Likert scale metrics for the level of detail/clarity of the sketched solutions, as this presents a positive correlation with their overall quality (Linsey et al., 2011). As shown in Linsey's work, the quality of the product solution increases as details are added. The choice of a 4-level scale enables rating the sketches with a sufficient degree of granularity, distinguishing from those barely understandable (lowest rating) to those perfectly clear and detailed (highest). The other two intermediate levels distinguish sketches that are somehow understandable but



https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

not fully clear (low-intermediate) and those which are clear and enriched with few descriptions (intermediate-high). This even-level scale of evaluation, different from the 3-level scale by Linsey, makes it possible to group the results into two big clusters without any need of disambiguating the ones rated at an intermediate level. The two experts reached an agreement above 80% after the independent rating. The complete agreement was achieved through a meeting to reach a consensus.

### 3.3 EEG data processing

To study the EEG characteristics across different bands, coherently with the Nyquist-Shannon sampling theorem, the sampling frequency of a minimum of 128 Hz enables the analysis of data bands up to the lower gamma-band [30-45] Hz. Among the frequency bands lower than the gamma-band, filtered from the data acquired using the commercial EEG headset, it is noticeable that the delta-band [0.1- 4] Hz gets highly contaminated and is hence no longer considered in the current research.

Figure 3 shows the signal processing pipeline adopted in the present paper. The efficacy of prevalent component analysis algorithms, namely ICA, PCA, and CCA, in extracting power is contingent upon the nature and quantity of artefacts present in a given dataset. Due to the uneven distribution and unbalanced quantity of artefacts in the design stages of the present study, these data treatment algorithms are no longer suitable. Thus, the present pipeline solely incorporates the preliminary band-pass filters and the logical distinction of artefacts referenced individually.

The EEG data were recorded throughout the entire experimental activity for each participant, segmented based on the duration of each task/stage. For this experiment, the task is DwMT and the stages are as in section 3.1. The Infinite Impulse Response (IIR) filter processes each segment of the raw data to remove the DC offset, typical of the headset used for the experiment. Then, a band-pass filter cuts off frequencies outside 4-45 Hz, coherent with what has been stated above. The data are further filtered into sub-frequency bands (theta, alpha, beta, gamma). Artefacts in the data are typically one order of magnitude -or more- bigger than brainwave data; their removal is necessary to generate a clean task dataset for the analysis. A subject-based threshold to exclude such outliers is defined by means of their baseline. The analysis of the EEG tracks enabled the visual association of artefacts to body movements video-recorded as for the protocol and the threshold is set accordingly. Its value corresponds to 16 times the task power (POW), which reflects 4 times the amplitude of the EEG signal. It is estimated twice in order to remove stronger artefacts first, then wipe out less intense ones. After the first round (dashed line in Fig.3), the baseline data is cleaned for the first time with a moving window (length: 0.25s; shift: 25%) to remove windows whose median POW exceeds the threshold. The threshold is again calculated (two dots-dashed lines) using this partially cleaned baseline. The new threshold aims at excluding outliers from stage-based data and generating band-based EED data by subject. Eventually, the median POW of each dataset was calculated to obtain the Task Related Power (TRP) according to the following formula.

$$TRP_{ij} = \frac{POW_{ij}(task_n)}{POW_{ij}(task_m)} \tag{1}$$

The task-based formula can also be used for stage-based analysis. The POW calculated through the pipeline for each participant $j$ has one data value for each electrode $i$ at each stage. Setting the denominator from the stage earlier than the stage in the numerator, the value of the TRP would facilitate the visualisation of brain activation dynamics. From stage $m$ to stage $n$, a TRP above 1 indicates a task-related power increase (TRPI), while a TRP below 1 implies a decrease (TRPD). The two clusters of participants are thereout formulated for the following analysis.

The whole data processing pipeline leverages an original Matlab script and the EEGLab toolbox (Delorme and Makeig, 2004)

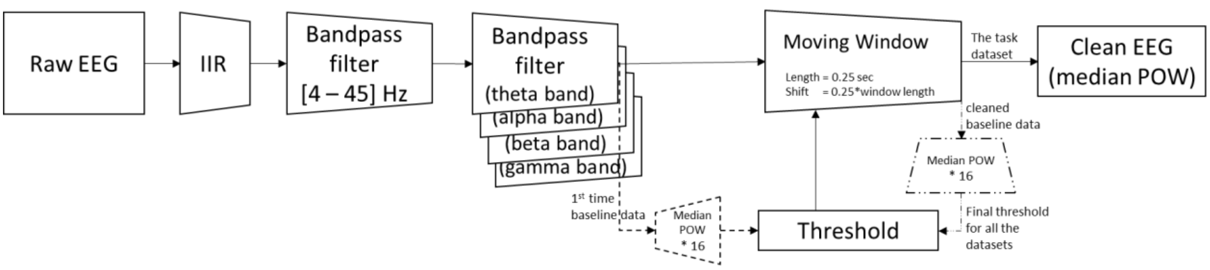



*Figure 3. The signal-processing pipeline*



https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

### 3.4 Criteria for group clustering the statistical review and the performance estimation

Carrying the goal of seeking all the relevant frequency bands and brain areas (channels) for distinguishing performance quality using EEG variation, the statistical tests are arrayed, as shown in Figure 4. Initially, the pipeline takes one specific sub-frequency band *X* and one stage comparison between stage *n* against stage *m*, and tabulates the TRP values of all the channels from all *N* participants. Then, it generates the full set of possible channel combinations, calculates the average TRP of the combined channels and associates the two values, for all *N* subjects. Each subject's dataset is also associated with the estimation of its design performance, as for the categorical metrics of Section 3.2. After that, two groups, TRPI and TRPD, are formulated accordingly.

After defining the two groups, a 2-sample t-test is used to confirm/reject the null hypothesis, which assumes the TRP values in two groups come from independent random samples from the normal distributions with equal means and equal but unknown variance. Bonferroni correction was applied using the total number of tests to keep the overall statistical significance below 0.05. The combinations of channels that result in rejecting the null hypothesis are further processed with the Kruskal-Wallis test on data about design performance. This checks whether the null hypothesis that the scores of the clarity assigned to subjects belonging to TRPI or TRPD come from the same distribution. The rejection of the null hypothesis indicates that the TRP measured with that channel combination within the specific frequency band allows the distinction of design performance quality. The threshold for significance for both tests is $p \leq 0.05$. Both the frequency and the channel combinations, in addition to the TRPI/TRPD performance distinction, are stored for further analysis.

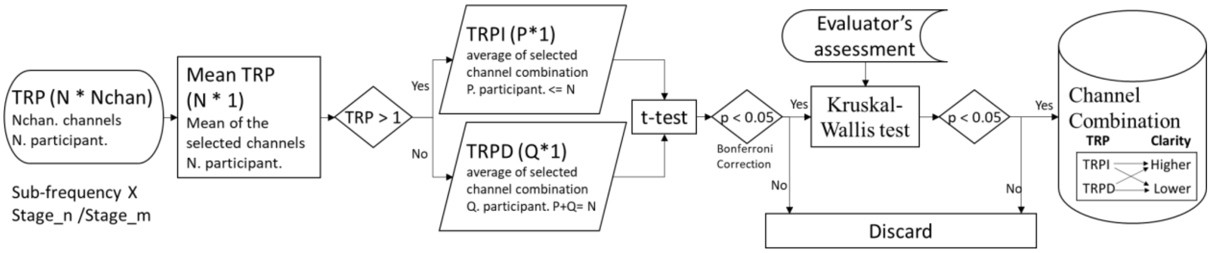



*Figure 4. The pipeline for clustering participants and statistical tests*

### 3.5 Criteria and the quality of performance estimation

The statistical tests enable filtering the combinations of the channels into the subsets, which could distinguish the quality of the outcomes between the two groups based on the EEG variation and the confusion matrix assesses how much the TRPI/TRPD model estimates design outcome quality.

These results help associate EEG variation with the outcome quality as a preliminary step towards the establishment of a predictive model. This paper focuses only on the EEG variations from the analysing and selecting stage (Stage 2) to the sketching stage (Stage 3). The expected model is meant to correlate TRPD to higher quality (TRPD-HQ) of the design outcome with the same qualitative evaluation carried out by the evaluators. The confusion matrix is built accordingly (Table 1).

*Table 1. Confusion matrix*

| | Raters - Higher Quality (HQ) | Raters - Lower Quality (LQ) |
|---|---|---|
| TRPD – Higher Quality (HQ) | TP | FP |
| TRPI – Lower Quality (LQ) | FN | TN |

The true positive (TP) derives when the data shows TRPD and therefore it is predicted to be higher quality, which is congruent with the evaluator's assessment. Or false positive (FP) when such prediction is opposite from the score given by the evaluator. Then when the data shows TRPI, which predicts a lower quality is congruent with the evaluator's result, it is a true negative (TN). Or it is a false negative (FN) when they are different. Then we selected the most frequently used performance metrics for classification based on these values, which are *accuracy* (2a), *precision* (2b), and *recall* (2c).

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN}; \qquad Precision = \frac{TP}{TP+FP}; \qquad Recall = \frac{TP}{TP+FN} \qquad (2a, 2b, 2c)$$



https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

*Accuracy* summarises the correct prediction from TRPD to higher quality and TRPI to lower quality. The *precision* denotes the proportion of participants who presented TRPD across the two design stages whose outcomes were deemed of higher quality by the evaluators. Last, the *recall* metric enables understanding how many participants the evaluators rated as the ones producing outcomes of higher quality also had TRPD across the stages.

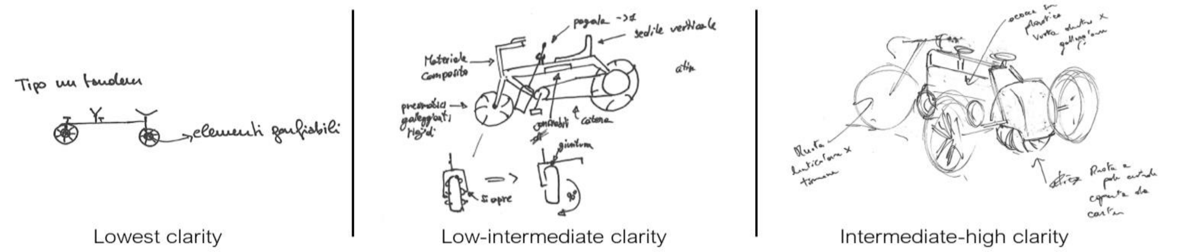
# 4 PARTICIPANTS AND EEG ACQUISITION DEVICE CHARACTERISTICS

Overall, 37 volunteers with a background in the engineering field were recruited for the experiment. Among these, 4 EEG datasets had to be discarded due to incomplete recording caused by the program crash that occurred during the execution of the DwMT task. Eventually, these 33 samples comprise 1 English speaker (Spanish male, age = 28), 7 Chinese native speakers (6 female, age M=29.33, SD=5.01, 1 male, age=30), and 25 Italian speakers (6 females, age M=23.17, SD=4.92, 19 males, age M=24.95, SD= 4.94). All the subjects received the task description in a language they can think in. At least two native speakers did the translation of the task description for each language. Thirty participants claimed to be right-hand dominant, while the other 3 have no side-of-hand dominance but wrote with their right hands only. None of them has any neurological deficits.
The headset used for EEG data collection is Emotiv EPOC X. It has a pre-mounted frame with 14 electrodes (AF3/4, F7/8, F3/4, FC5/6, T7/8, P7/8, O1/2), sampling at 128Hz. The raw data is accessible to run signal processing according to the needs.

# 5 RESULT

## 5.1 Evaluation of the design outcome

The two evaluators made the assessment independently on all 37 collected sketches according to the criteria described in section 3.2 and then met for consensus on the 4 sketches on which they did not agree. Then, the 4 subjects without EEG datasets were excluded from further analysis. Figure 5 shows samples of the rated sketches. Limited by the experience of the participants and the number of samples, around half were evaluated as having the lowest clarity by both evaluators, and no one received the highest clarity score. Hence the final groups' division differentiates only between the lowest level and the level above. Finally, both evaluators agreed and assigned 15 outcomes to the lower-quality group, and the rest 18 to the higher-quality group.




Lowest clarity | Low-intermediate clarity | Intermediate-high clarity

*Figure 5. Samples of outcomes*

## 5.2 Statistical analysis and estimation results

The statistical tests check for significant group distribution differences through all possible combinations of the 14 channels (16383 different combinations) for each of the four sub-frequency bands and also for the total bandwidth and then against the group division made by the evaluator. The results returned from the tests suggested a significant link between the TRPD and the higher quality of the outcome, using the assessment result from the evaluators (245 cases of TRPD-HQ against 6 cases of TRPI-HQ). The summarised results of the TRPD-HQ cases in which performance metrics yield above 0.7 are presented below in Table 2.
All the cases returned from the statistical tests show an *accuracy* value greater than or equal to 0.7, among which the maximum *accuracy* reaches 0.727. While the number of cases was lowered to 35 to achieve a *precision* from 0.7 to 0.765. And 228 cases yield values from 0.7 to 0.889 for the *recall*.
The theta band is the most informative for predicting design outcome quality as it returns higher values of *accuracy* and *recall* in the majority of cases (214 for both *precision* and *recall*), while the rest are in the alpha band. This means that both theta and alpha bands are relevant. However, to



https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

guarantee higher overall correctness of estimation or that fewer results are falsely predicted as LQ by observing TRPI, the theta band appears to be the most promising. In order to improve *precision* by reducing false positive cases (TRPD, but assessed as LQ by the evaluator), alpha is the only frequency band that the model suggests considering.

Considering the essential channels to be included in the model, the channels are ranked by the number of occurrences to have at least 70% of the significant cases under each metric. At least 5 channels are suggested to be included in combination with higher priority to obtain higher *accuracy* and *recall*. The top 5 used channels are listed for these two metrics. A significant overlapping could be observed at the top 4 used channels, namely AF3, F3, FC5, and AF4, all in the frontal area, mainly on the left side. *Accuracy* and *recall* have also two different channels that might work as a proxy for the quality of the solution: O1 and FC6, respectively. This means the left occipital area also plays a significant role in the *accuracy* of the estimation. At least 3 channels should be included to reach higher *precision*, and the three most used channels are O1, F7 and F4. It covers areas from frontal to occipital sites. The highest and the lowest percentage of the occurrence are also shown in Table 2.

*Table 2. Summarisation of the significant results, for each metric's value above 0.7, using TRPD between stages to predict the higher quality of clarity*

| Metrics (>= 0.7) | Max Value | Frequency Band | No. Channels | Most used Channels (ranked by N. occurrence%) |
|---|---|---|---|---|
| *Accuracy* (245 cases) | 0.727 | θ or α | 5 - 8 | AF3(87%), F3, FC5, AF4, O1 (64%) |
| *Precision* (35 cases) | 0.765 | α | 3 - 6 | O1 (80%), F7, F4(63%) |
| *Recall* (228 cases) | 0.889 | θ or α | 5 - 8 | AF3(89%), F3, FC5, AF4, FC6 (64%) |

## 6 DISCUSSION

Our experimental results highlight there is a potential to use EEG (de) synchronisation to predict the quality of the design outcome of the specific tested design task across its stages. It appears that from the design stage of information perception and analysis to the stage of idea illustration, for the cohort unable to detail their ideas, a higher EEG power is more likely to be observed in the latter stage. But this correlation depends on the brain area (channel) and the frequency band.

The *accuracy* reaches higher values using data from theta and alpha bands, channels located primarily on the left frontal and left occipital area. These areas are usually associated with the working load during retrieval (Klimesch et al., 2006) and complex visual processing, which are embedded activities in both stages. The *accuracy* metric based on the TRPD-HQ model is up to 0.727 (24/33). Unfortunately, in the literature, there is no similar TRPI/TRPD-based investigation to compare the prediction model's performance. However, similar decision-making models (e.g. based on ERS/ERD) show relatively lower values for *accuracy* (Seeland et al., 2015). The model's performance might suffer from the difficulty that occurred when the evaluators tried to reach an agreement on the judgements, but it is also affected by unexpected huge behavioural differences across the two stages that happened to several participants. As for the design task, the two stages entailed both arm and neck movements. However, the given paper space couldn't explicitly restrict the dimension of the final sketch. For those who sketched larger images or those with certain sketching habits (i.e., swing the hand to draw lines, hash parts of the component), their EEG data are more contaminated by these body movements. The headset, beyond real data and the contribution due to myoelectricity, is also sensitive to these movements, as they could also induce sensor contact issues, adding artefacts, especially in the lower frequency bands with higher amplitude than the sole cognition activity. The result of observing TRPD indicates an effective artefact removal, considering that a larger portion of body movements occurred during sketching than during the analysing stage. On the contrary, TRPI should be prevalent if the treatment fails. Since, from the outcomes collected in our experiment, there is no direct correlation between the dimension of the sketches and their clarity, further confirmation regarding the impact of body movements would help us better understand to what extent we should limit the interface in future studies.

The *precision* metric's value suggests that the alpha band from the left occipital and frontal areas might serve as a good proxy for objectively assessing the quality of design outcomes, still within the TRPD model.


https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

The *recall* for the TRPD model appears to reach the highest value among the three performance metrics. The result indicates when the evaluators couldn't reach an agreement in finding higher quality outcomes, EEG data acquired from channels in the left frontal area at theta band might assist with suggestions.

Clinical and psychological studies have already reported that the variations in the alpha band power are sensitive to factors such as attentional demands (Ray and Cole, 1985), use of working memory (Stipacek et al., 2003), and task difficulty (Fink et al., 2005). Greater subjects' cognitive demands correspond to a more significant power decrease. These findings also provide insights into what may be the hidden reasons behind the current findings in design neurocognition. The participant's attitude towards different design stages could lead to different quality of the design outcome. One might pay more attention to the presentation of the final idea than to the careful selection of the elements and therefore receive a higher score for clarity. Or one might pay more attention to analysing the problem and making choices than to illustrating the idea. In both situations, the EEG variation could prompt the designer to balance the effort and rearrange the activities to achieve a better outcome, reducing iterations and slowdowns in the design process, thus saving the designer's time and helping the transition from ex-post evaluation to real-time (objective) monitoring.

## 7 CONCLUSION AND FUTURE DEVELOPMENT

In conclusion, the current study proves that the quality of the design outcome, in terms of the level of detail/clarity of the generated sketches, can be correlated to EEG variations. The model that links EEG task-related power desynchronisation to high-quality outcomes showed that some channel combinations, in specific frequency bands, have a strong association with the clarity of the generated sketches. This might help the assessment of ideas-as-sketches more objectively and, in the future, might enable real-time biofeedback to the designer.

The reported results are based on one specific design task, including only two stages in the whole design process. Further verification could include the first reading stage at the same design task to complete the understanding of the dynamic of the brain activation across the entire design task with the morphological table. The two analysed stages (selecting vs sketching ideas) both involve activities that are essential to a general design task. The possibility of extending the current study to other design protocols is not restricted. The samples are collected from diverse cultural backgrounds and are imbalanced in gender and age. An expansion of the sample is required to confirm this initial evidence and for more comprehensive result validation.

## REFERENCES


Adolphy, S., Gericke, K. and Blessing, L. (2009), “Estimation and its Role in Engineering Design - An Introduction”, Proceedings of the 17th International Conference on Engineering Design, pp. 255–266.

Benedek, M., Bergner, S., Könen, T., Fink, A. and Neubauer, A.C. (2011), “EEG alpha synchronization is related to top-down processing in convergent and divergent thinking”, Neuropsychologia, Pergamon, Vol. 49 No. 12, pp. 3505–3511, https://doi.org/10.1016/j.neuropsychologia.2011.09.004.

Delorme, A. and Makeig, S. (2004), “EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis”, Journal of Neuroscience Methods, Elsevier, Vol. 134 No. 1, pp. 9–21, https://doi.org/10.1016/j.jneumeth.2003.10.009.

Fink, A., Grabner, R.H., Neuper, C. and Neubauer, A.C. (2005), “EEG alpha band dissociation with increasing task demands”, Cognitive Brain Research, Vol. 24 No. 2, pp. 252–259, https://doi.org/10.1016/j.cogbrainres.2005.02.002.

Fink, A. and Neubauer, A.C. (2006), “EEG alpha oscillations during the performance of verbal creativity tasks: Differential effects of sex and verbal intelligence”, International Journal of Psychophysiology, Vol. 62 No. 1, pp. 46–53, https://doi.org/10.1016/j.ijpsycho.2006.01.001.

Foldes, S.T. and Taylor, D.M. (2013), “Speaking and cognitive distractions during EEG-based brain control of a virtual neuroprosthesis-arm”, Journal of NeuroEngineering and Rehabilitation, Vol. 10 No. 1, pp. 3–6, https://doi.org/10.1186/1743-0003-10-116.

Hu, Y., Ouyang, J., Wang, H., Zhang, J., Liu, A., Min, X. and Du, X. (2022), “Design Meets Neuroscience: An Electroencephalogram Study of Design Thinking in Concept Generation Phase”, Frontiers in Psychology.

Hummel, F.C. and Gerloff, C. (2006), “Chapter 15 Interregional long-range and short-range synchrony: a basis for complex sensorimotor processing”, Progress in Brain Research, Vol. 159 No. 06, pp. 223–236, https://doi.org/10.1016/S0079-6123(06)59015-6.




https://doi.org/10.1017/pds.2023.154 Published online by Cambridge University Press

Jaarsveld, S., Fink, A., Rinner, M., Schwab, D., Benedek, M. and Lachmann, T. (2015), “Intelligence in creative processes: An EEG study”, Intelligence, Elsevier Inc., Vol. 49, pp. 171–178, https://doi.org/10.1016/j.intell.2015.01.012.

Jansson, D.G. and Smith, S.M. (1991), “Design fixation”, Design Studies, Vol. 12 No. 1, pp. 3–11, https://doi.org/10.1016/0142-694X(91)90003-F.

Jia, W. and Zeng, Y. (2021), “EEG signals respond differently to idea generation, idea evolution and evaluation in a loosely controlled creativity experiment”, Scientific Reports, Nature Publishing Group UK, Vol. 11 No. 1, pp. 1–20, https://doi.org/10.1038/s41598-021-81655-0.

Klimesch, W., Doppelmayr, M. and Hanslmayr, S. (2006), “Chapter 10 Upper alpha ERD and absolute power: their meaning for memory performance”, Progress in Brain Research, Vol. 159, pp. 151–165, https://doi.org/10.1016/S0079-6123(06)59010-7.

Kulkarni, A., Chong, D. and Batarseh, F.A. (2020), “Foundations of data imbalance and solutions for a data democracy”, Data Democracy, Elsevier, pp. 83–106, https://doi.org/10.1016/B978-0-12-818366-3.00005-8.

Lan, Z., Sourina, O., Wang, L. and Liu, Y. (2016), “Real-time EEG-based emotion monitoring using stable features”, Visual Computer, Springer Berlin Heidelberg, Vol. 32 No. 3, pp. 347–358, https://doi.org/10.1007/s00371-015-1183-y.

Li, S., Becattini, N. and Cascini, G. (2021), “Correlating Design Performance To Eeg Activation: Early Evidence From Experimental Data”, Proceedings of the Design Society, Vol. 1 No. AUGUST, pp. 771–780, https://doi.org/10.1017/pds.2021.77.

Linsey, J.S., Clauss, E.F., Kurtoglu, T., Murphy, J.T., Wood, K.L. and Markman, A.B. (2011), “An Experimental Study of Group Idea Generation Techniques: Understanding the Roles of Idea Representation and Viewing Methods”, Journal of Mechanical Design, Vol. 133 No. 3, https://doi.org/10.1115/1.4003498.

Liu, Y.-C., Chang, C.-C., Yang, Y.-H.S. and Liang, C. (2018), “Spontaneous analogising caused by text stimuli in design thinking: differences between higher- and lower-creativity groups”, Cognitive Neurodynamics, Vol. 12 No. 1, pp. 55–71, https://doi.org/10.1007/s11571-017-9454-0.

Lopes da Silva, F. (1991), “Neural mechanisms underlying brain waves: from neural membranes to networks”, Electroencephalography and Clinical Neurophysiology, Vol. 79 No. 2, pp. 81–93, https://doi.org/10.1016/0013-4694(91)90044-5.

Lopes da Silva, F.H. (2006), “Chapter 1 Event-related neural activities: what about phase?”, Progress in Brain Research, Vol. 159, pp. 3–17, https://doi.org/10.1016/S0079-6123(06)59001-6.

Lukacevic, F., Li, S., Becattini, N. and Škec, S. (2022), “Comparing EEG Brain Power of Mechanical Engineers in 3D CAD Modelling from 2D and 3D Representations”, Proceedings of the Design Society, Vol. 2 No. May, pp. 901–910, https://doi.org/10.1017/pds.2022.92.

Peirce, J., Hirst, R. and MacAskill, M. (2022), Building Experiments in PsychoPy, 2nd ed., SAGE Publications Inc, London.

Pfurtscheller, G. and Lopes Da Silva, F.H. (1999), “Event-related EEG/MEG synchronization and desynchronization: Basic principles”, Clinical Neurophysiology, Vol. 110 No. 11, pp. 1842–1857, https://doi.org/10.1016/S1388-2457(99)00141-8.

Pfurtscheller, G., Neuper, C. and Kalcher, J. (1993), “40-Hz oscillations during motor behavior in man”, Neuroscience Letters, Vol. 164 No. 1–2, pp. 179–182, https://doi.org/10.1016/0304-3940(93)90886-P.

Pfurtscheller, G. and Silva, F.L. da. (2017), EEG Event-Related Desynchronization and Event-Related Synchronization, edited by Schomer, D.L. and Lopes da Silva, F.H., Vol. 1, Oxford University Press, https://doi.org/10.1093/med/9780190228484.003.0040.

Ratcliff, R., Philiastides, M.G. and Sajda, P. (2009), “Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG”, Proceedings of the National Academy of Sciences, Vol. 106 No. 16, pp. 6539–6544, https://doi.org/10.1073/pnas.0812589106.

Ray, W.J. and Cole, H.W. (1985), “EEG Alpha Activity Reflects Attentional Demands, and Beta Activity Reflects Emotional and Cognitive Processes”, Science, Vol. 228 No. 4700, pp. 750–752, https://doi.org/10.1126/science.3992243.

Seeland, A., Manca, L., Kirchner, F. and Kirchner, E.A. (2015), “Spatio-temporal Comparison between ERD/ERS and MRCP-based Movement Prediction”, Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, SCITEPRESS - Science and Technology Publications, pp. 219–226, https://doi.org/10.5220/0005214002190226.

Stern, J.M. (2013), Atlas of EEG Patterns, 2nd ed., Lippincott Williams and Wilkins., Philadelphia.

Stipacek, A., Grabner, R.H., Neuper, C., Fink, A. and Neubauer, A.C. (2003), “Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load”, Neuroscience Letters, Vol. 353 No. 3, pp. 193–196, https://doi.org/10.1016/j.neulet.2003.09.044.

Vieira, S., Benedek, M., Gero, J., Li, S. and Cascini, G. (2022), “Design spaces and EEG frequency band power in constrained and open design”, International Journal of Design Creativity and Innovation, Taylor & Francis, Vol. 10 No. 4, pp. 193–221, https://doi.org/10.1080/21650349.2022.2048697.