



OPEN

DATA DESCRIPTOR

# EEG-based BCI Dataset of Semantic Concepts for Imagination and Perception Tasks

Holly Wilson<sup>1</sup>✉, Mohammad Golbabaee<sup>2</sup>, Michael J. Proulx<sup>3</sup>, Stephen Charles<sup>1</sup>  
& Eamonn O'Neill<sup>1</sup>✉

Electroencephalography (EEG) is a widely-used neuroimaging technique in Brain Computer Interfaces (BCIs) due to its non-invasive nature, accessibility and high temporal resolution. A range of input representations has been explored for BCIs. The same semantic meaning can be conveyed in different representations, such as visual (orthographic and pictorial) and auditory (spoken words). These stimuli representations can be either imagined or perceived by the BCI user. In particular, there is a scarcity of existing open source EEG datasets for imagined visual content, and to our knowledge there are no open source EEG datasets for semantics captured through multiple sensory modalities for both perceived and imagined content. Here we present an open source multisensory imagination and perception dataset, with twelve participants, acquired with a 124 EEG channel system. The aim is for the dataset to be open for purposes such as BCI related decoding and for better understanding the neural mechanisms behind perception, imagination and across the sensory modalities when the semantic category is held constant.

## Background & Summary

Brain computer interfacing and cognitive neuroscience are fields which rely on high quality brain activity based datasets. Surface electroencephalography (EEG) is a popular choice of neuroimaging technique for BCIs due to its accessibility in terms of cost and mobility, its high temporal resolution and non-invasiveness. Although EEG datasets can be time consuming and expensive to obtain, they are extremely valuable. A single open source dataset can form the basis of many varied research projects, and thus can more rapidly advance scientific progress. For example, EEG datasets for inner speech commands<sup>1</sup> and for object recognition<sup>2</sup> were recently created and shared to address a lack of publicly available datasets in these areas. These datasets enable the development of sophisticated techniques for analysis and decoding, which can be used to investigate neural representation mechanisms and improve decoding performance for EEG based BCIs.

Different paradigms have been used for EEG based BCIs such as Event Related Potential (ERP) BCIs for decoding inner speech<sup>1,3</sup>, Steady-State Visual Evoked Potentials (SSVEPs)<sup>4</sup> and motor imagery<sup>5</sup>, and oscillatory activity driven BCIs for tasks such as drowsiness detection<sup>6</sup>. Recently, there has been growing interest in decoding alternative information forms such as auditory and visual, perception and imagination<sup>7</sup>, and semantic information<sup>8</sup>. However, the lack of open source EEG datasets for decoding imagined and perceived semantic level information is hindering progress towards this research goal.

Visual decoding involves decoding simple low level visual components such as colour and shape, or complex naturalistic images of objects, scenes and faces. In contrast, semantic decoding extracts conceptual information such as object types or classes. For example, was the object in an image shown to an observer a flower or a guitar? The low level visual and auditory sensory details of the semantic concept, such as whether the flower is yellow or purple, are ignored with a focus on the high level meaning of 'flower'. The advantage of decoding semantic information, as opposed to sensory based information such as visual details, is that semantic representation is partially invariant across modalities<sup>9-13</sup>. Invariance to low level sensory detail can be considered a desirable quality in BCI systems in which within class generalisability is a key goal. This can help increase robustness to real world data heterogeneity.

<sup>1</sup>Department of Computer Science, University of Bath, Bath, BA2 7AY, UK. <sup>2</sup>Department of Engineering Mathematics, University of Bristol, Bristol, BS8 1TW, UK. <sup>3</sup>Department of Psychology, University of Bath, Bath, BA2 7AY, UK. ✉e-mail: [hlw69@bath.ac.uk](mailto:hlw69@bath.ac.uk); [maseon@bath.ac.uk](mailto:maseon@bath.ac.uk)

| Metric | 3_3  | 8_3 | 10_1 | 11_1 | 12_1 | 12_2 | 13_1 | 14_1 | 14_2 | 15_1 | 15_2 | 16_1 | 17_1 | 18_2 | 19_1 | Avg  | Std  |
|--------|------|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| VVIQ   | 4.25 | 3.1 | 4.9  | 3.3  | 3.2  | 3.2  | —    | 4.2  | 4.2  | 4.1  | 4.1  | —    | 3.5  | 3.6  | 3.1  | 3.75 | 0.55 |
| BAIS   | 4.8  | 4.2 | 6.33 | 4    | 3.8  | 3.8  | —    | 5    | 5    | 5.7  | 5.7  | —    | 4    | 5.6  | 3.8  | 4.76 | 0.85 |

**Table 1.** The Vividness of Visual Imagery Questionnaire (VVIQ) and Bucknell Auditory Imagery Scale (BAIS) scores for each participant, session and the average (avg) and standard deviation (std). For two participants, sub-13 and sub-16, the vividness questionnaires were not completed.

Growing evidence of neural overlap between perception and imagination<sup>14,15</sup> may also facilitate generalisability. This task invariance has enabled cross-decoding between perception and imagination<sup>16,17</sup>. Efforts are being made to determine the spatiotemporal extent of these shared neural representations<sup>18–20</sup>, which may be most invariant in brain regions and time points associated with latent representations; i.e. closer to semantic level information. For example, the differences between imagery and vision appear to be most pronounced in the early visual cortex, with greater overlap occurring higher up in the visual hierarchy<sup>21</sup> and at time points linked to high level perceptual processing<sup>14</sup>.

To drive decoding of a class based on semantic information, stimuli must vary in their low level sensory details. This approach was employed in a recent open source dataset that captured EEG measurements for object recognition using a rapid serial visual presentation paradigm<sup>2</sup>. The dataset includes 22,248 images related to 1,854 concepts. While there are impressive semantic decoding results emerging using fMRI<sup>22–25</sup> and EEG<sup>26–28</sup> which demonstrate feasibility, the field lacks an open source EEG dataset for researchers to investigate semantic representation across several sensory modalities, as well as both perception and imagination.

In this paper, we introduce a novel dataset, as well as the code for pre-processing and analysis, designed for investigating and decoding semantic representation of imagined and perceived visual and auditory information. We also present an initial analysis to demonstrate this dataset's utility. To capture semantic representation, we drive high variance within each class (or rather semantic category). Specifically, we use three semantic concepts—penguin, guitar and flower—that participants perceived and subsequently imagined in auditory, visual orthographic, and visual pictorial forms. Furthermore, we provide a metric for the vividness of imagination metric for each participant for both the visual and auditory modalities. Individual differences in imagination capacity are shown to impact neural correlates<sup>29,30</sup> and therefore may affect the decodability of, or the decoding strategy used for, each individual.

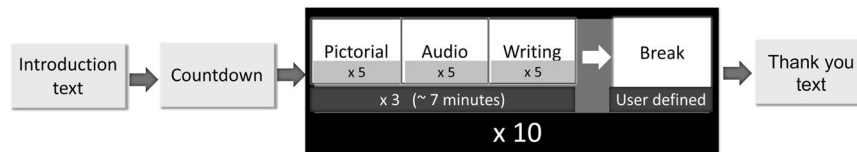
Some proposed uses of this dataset for both BCI and cognitive neuroscience oriented research questions include:

1. Decoding between sensory modalities such as auditory, visual orthographic and visual pictorial.
2. Decoding task type, specifically between perception and imagination.
3. Decoding the semantic category regardless of the sensory modality presentation or task.

## Methods

**Participants.** Ethics approval was obtained from the Psychology Research Ethics Committee at the University of Bath (Ethics code: 19–302). Participants gave informed consent to take part in this study and for their data to be shared. Eight participants were recruited for a data collection pilot to ensure the quality of the dataset. This allowed us to identify and address any syncing issues with the Lab Streaming Layer network, as well as unexpected environmental noise at around 27 Hz in two of the sessions. The final version of the experiment was completed by twelve participants, most of whom were students at the University of Bath. Initially, selection criteria included normal or corrected vision and hearing, and excluded individuals with epilepsy. However, we later expanded the criteria to include individuals with visual and hearing impairments, to enable our dataset to support a wider range of research questions. One participant with visual and hearing impairment was included in the final sample. Participants were reimbursed £20 for their time in exchange for participating in an approximately two hour session.

**Experimental procedure.** Participants were offered the opportunity to participate in a second data gathering session, in order to increase the number of trials for each participant. Of the twelve participants, nine completed one session and three returned for a second session. The experiment was conducted in a soundproof and lightproof room. It was not electrically shielded but all mains outlets other than the acquisition laptop charge point were turned off. The EEG setup, including cap fitting and gel application to the electrodes, took approximately 40 to 60 minutes. During the first session, participants completed two questionnaires while the gel was being applied: the vividness subscale of the Bucknell Auditory Imagery Scale (BAIS-V)<sup>31</sup> and the Vividness of Visual Imagery Questionnaire (VVIQ)<sup>32</sup>. Subsequent to this, participants performed a practice version of the experimental tasks with a chance to ask questions around any uncertainties. After the setup was complete, the light was turned off, the experimenters left the testing room and went into an adjacent room, and the participant began the study when ready by pressing the computer keyboard's space bar. For a schematic of the main task flow, see Fig. 1. The experiment was designed using Psychopy Version 3<sup>33</sup>, and presented on a 1920 × 1080 resolution screen. The Psychopy files are made available as described in the Usage Notes section. The ANT Neuro acquisition software 'eego' was used to record the EEG data. A Lab Streaming Layer (LSL) network sent the triggers from the presentation PC to the acquisition software to time-stamp the stimuli and task relevant information. There were



**Fig. 1** This figure shows an example of a pictorial trial. After a cue indicating whether the upcoming task is pictorial, orthographic or audio, five trials occur with a different stimulus used in each. Before the break, one block of each type of modality is cycled through, which takes around seven minutes. The duration of each break is chosen by the participant.

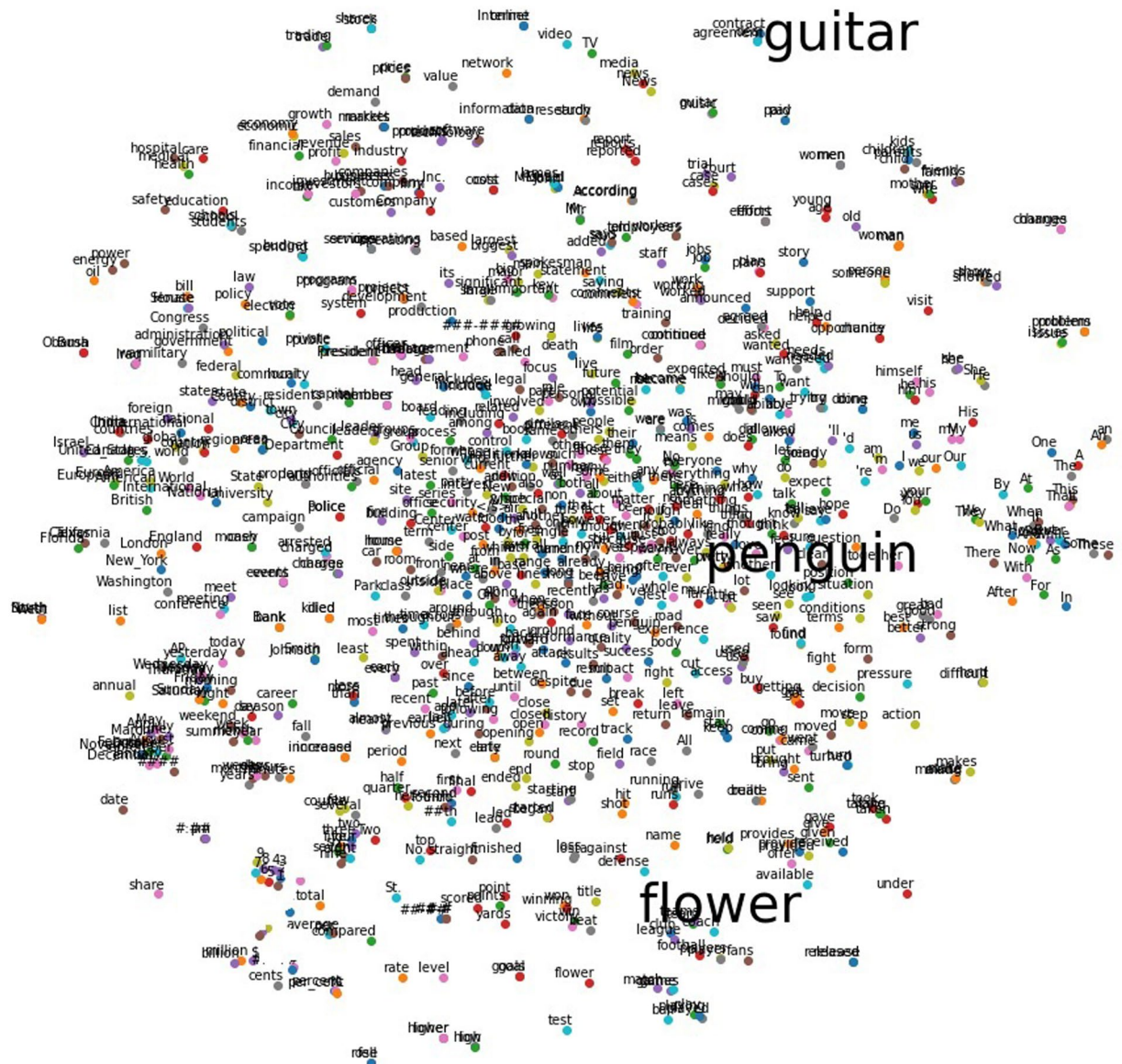
ten blocks in total, though the majority of participants did not complete all the blocks due to fatigue or reporting reduced concentration. See Table 4 for the amount of trials completed for each condition for each participant. The participants were encouraged to take breaks between each block and call the experimenter if they required water or had any concerns.

**Questionnaire.** The VVIQ and BAIS-V are self-report measures of mental imagery ability. The BAIS-V is a subscale of BAIS which captures the subjective clarity or vividness of an imagined sound, such as a trumpet playing happy birthday, on a scale of 1–7. A score of 7 is as vivid as the actual sound, whereas 1 indicates there was no sound at all. The VVIQ measures the subjective vividness of an imagined scenario such as a sunset, on a scale of 1–5, with 5 being the most vivid and 1 meaning no image at all. For VVIQ and BAIS-V results, see Table 1. The mean VVIQ score was 3.75 (std = 0.55) and average BAIS-V was 4.76 (std = 0.85). VVIQ and BAIS scores are significantly correlated as calculated using Spearman's Rank with  $r = 0.79$  and  $p = 0.007$ .

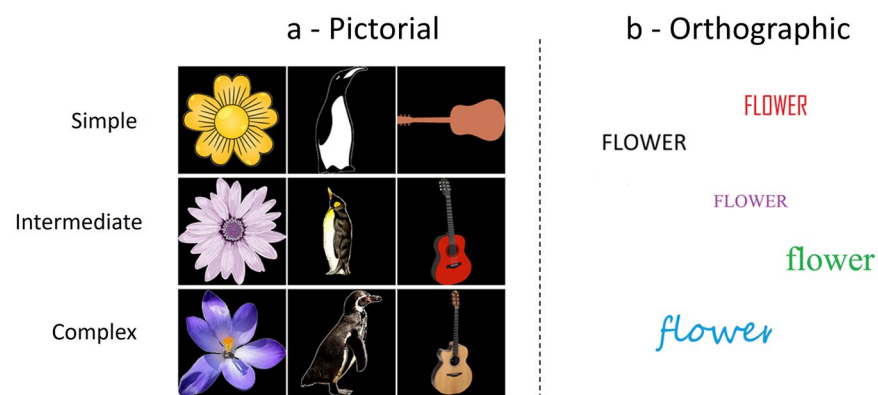
**Data acquisition.** A 128 channel *ANT Neuro eego Mylab* measuring system (ANT Neuro B.V., Hengelo, Netherlands) was used, with 124 EEG electrodes. The gel-based waveguard cap has active shielding which protects the signal from 50/60 Hz environmental noise. The sampling rate was 1024 Hz, with a 24-bit resolution. The montage, with pre-fixed electrode positions, is laid out according to the five percent electrode system<sup>34</sup>, which is an extension from the standard 10/20 layout for higher resolution EEG systems. The EEG cap size was selected based on the participant's head circumference in cm. Large is 56–61 cm, medium is 51–56 cm and small is 47–51 cm. Once the cap was fitted to the participant's head, OneStep Cleargel conductive gel was applied to the electrodes with CPz as reference, and the ground fixed to the left mastoid with Ten20 paste. Impedance of below 50 was sought, but due to variables such as hair thickness and other factors, there were often up to ten electrodes that had higher impedance. After the experiment was finished, the recording was stopped and the EEG data were stored as *.cnt* files, and the events as *.evt* files in ANT Neuro native format.

**The paradigms.** This study involved six paradigm variations, consisting of two tasks: imagination or perception, and three sensory modalities: visual pictorial, visual orthographic and auditory comprehension. The semantic categories used were flower, penguin and guitar. These three categories were selected based on semantic distance and syllable length. Semantic distance was determined by computing a Word2Vec latent space<sup>35</sup>, where each word is represented as a vector and the distance between vector pairs signifies the semantic similarity of two words. The distance between each of the pairs was calculated to ensure all pair-distances were  $< 0.2$ . A visual plot was then created using a t-distributed Stochastic Neighbour Embedding (t-SNE) which enables high dimensional data to be visualised in a 2D space (see Fig. 2). While common daily objects may be preferred as stimuli for BCI purposes, we selected more obscure objects which are unlikely to be used in the same contexts. This decision was driven by two main factors. First, using objects that people encounter on a daily basis can introduce unpredictable semantic associations and relations from their daily routines. Secondly, objects we have expertise in processing, such as faces, may result in spatially clustered selectivity or brain modularity<sup>36</sup>. This can restrict the generalisability of findings to non-expertise categories and thereby reduce the overall scope of application. Another constraint in selecting the semantic categories was that they all have two syllables. It is crucial to keep syllable length constant in the auditory comprehension paradigm to ensure that decoding is based on semantic properties rather than the syllable number associated with different words.

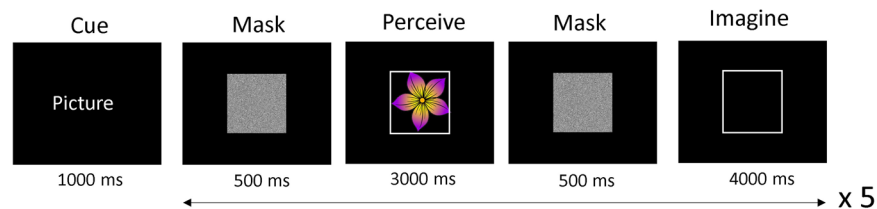
**Visual pictorial.** The visual pictorial paradigm involves perception and imagination of images belonging to the three semantic categories: flower, penguin and guitar. The visual pictorial stimuli consisted of coloured images with a resolution of  $1200 \times 1200$  pixels against a black background (see Figs. 3, 4). In the context of object representation, incorporating objects within a consistent scene can enhance their semantic relations and aid in their recognition<sup>37</sup>. However, to maintain the purity of our study's semantic concepts we opted to exclude any contextual scene information. This was because the addition of contextual information could potentially introduce unexpected semantic associations, thus introducing semantic noise. Furthermore, including contextual scenes would have added complexity, making the imagination task more challenging and potentially leading to increased participant fatigue. Therefore, we chose to focus solely on the objects themselves, without any accompanying contextual information. There are three levels of complexity for the images: simple, intermediate and naturalistic. For both flowers and guitars, there are eight different exemplars for the simple level and nine each of the intermediate and naturalistic levels. For penguin, there are nine exemplars for each level of complexity.



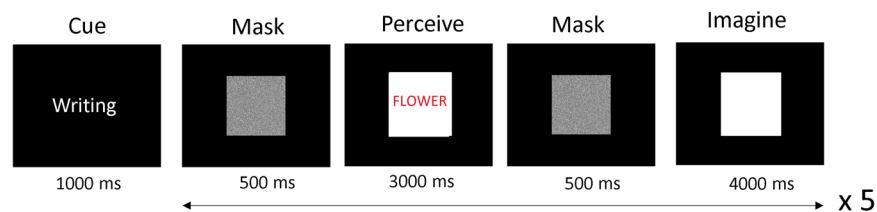
**Fig. 2** Visualisation demonstrating that the three selected semantic words (penguin, guitar and flower) are semantically distant from each other. The distances, computed using Word2Vec, are plotted in 2D using t-SNE.



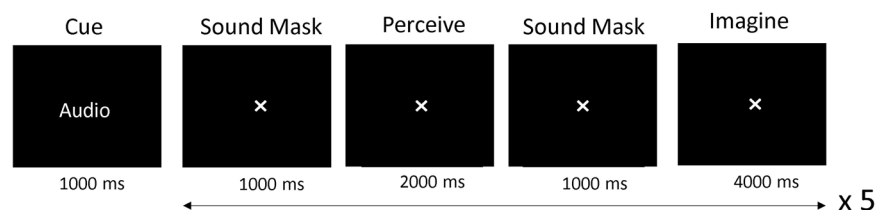
**Fig. 3** Examples of the visual (a) pictorial and (b) orthographic stimuli used in the experiment. Pictorial stimuli ranged in complexity from simple to intermediate to naturalistic, while orthographic stimuli varied in colour and font.



**Fig. 4** An example of a pictorial trial. After the cue, 5 trials occur with a different picture used in each. The picture is bounded in a white box, which reappears to frame the mental image for the imagination trial.



**Fig. 5** Example of an orthographic trial. After the cue, 5 trials occur with a different orthographic representation used in each. The written word appears against a white background, which reappears in the imagination trial to ensure similar scaling between imagination and perception.



**Fig. 6** Example of an auditory trial. After the cue, 5 trials occur with a different spoken word recording used in each. A white noise sound mask of 1000 ms is used to prevent residual stimulus audio representation leaking between the perception and imagination trials.

**Visual orthographic.** Visual orthographic is the image of the word form of the semantic categories. The stimuli consisted of a  $1200 \times 1200$  pixel white background with writing of either ‘penguin’, ‘flower’ or ‘guitar’ overlaid (see Figs. 3, 5). There were 30 exemplars for each category, with five different colours used (black, blue, red, green, purple) and six different font styles.

**Auditory comprehension.** Auditory comprehension consists of the speech version of the three semantic categories ‘penguin’, ‘guitar’ and ‘flower’. Recordings of these words were obtained from different speakers who did not participate in the EEG experiment. In the perception task, participants passively listened to these recordings which were processed using Audacity to remove background noise. Each clip was two seconds long. The words were spoken in either a normal, low or high voice. During the imagination task, participants were asked to imagine the spoken words that they had heard, using the same voice of the speaker rather than their own inner voice. To view an example of an audio trial, refer to Fig. 6.

**Data processing. Bad channels.** To rigorously adjust for bad channels, a combination of manual and automatic bad channel detection was used. Bad channels identified from visual inspection of the plotted raw data in the ANT Neuro eego software were recorded in the *meta\_extended.csv* file, discussed in the Data Records section, for each participant and session. Automatic bad channel detection was computed using PyPrep PrepPipeline <https://pyprep.readthedocs.io/en/latest/generated/pyprep.PrepPipeline.html>. This method utilises several bad channel detection methods, including identifying channels that do not correlate with other channels, channels with abnormally low or high amplitudes, or high quantities of high frequency noise, and channels with flat signals. Channels were re-referenced before interpolation was applied to correct for bad channels.

**Re-referencing.** During acquisition, electrodes were referenced to CPz. Re-referencing was conducted after all steps that offset the statistical trend of the overall data. Re-referencing was applied before and then after bad channel interpolation using common average referencing in MNE. A third re-referencing step was applied after filtering to remove low frequency drifts.

| Task type   | Sensory Modality | Duration |
|-------------|------------------|----------|
| Perception  | Auditory         | 2 s      |
| Imagination | Auditory         | 4 s      |
| Perception  | Orthographic     | 3 s      |
| Imagination | Orthographic     | 4 s      |
| Perception  | Pictorial        | 3 s      |
| Imagination | Pictorial        | 4 s      |

**Table 2.** The duration in seconds for each type of epoch.

**Filtering.** Data were filtered to remove power-line noise via notch filtering. Powerline noise in the UK where this dataset was recorded is at 50 Hz, therefore we filter for 50 Hz and its harmonics: 100 and 150 Hz. We also remove low frequency drifts which arise from movements of the head, scalp perspiration and wires. Filtering out frequencies below 2 Hz, via high pass filters, is recommended for high quality ICA decompositions<sup>38</sup>.

**Artefact removal.** Artefacts include eye movements such as blinks and horizontal eye movements as well as muscle activity. Independent Component Analysis (ICA) was applied to the raw pre-processed data rather than epoched data. The FastICA algorithm was used, and 50 components selected. To identify eye components, we used an MNE implementation to generate epochs around electrooculogram (EOG) artefact events. These were estimated from channels close to the eyes ‘Fp1’ and ‘Fp2’. By estimating these artefacts, the components can then be rejected from the ICA components. The resulting data after ICA retains all 124 original dimensions.

**Epoching.** Event labels for each condition were used to identify the beginning of each epoch. As the `mne.Epochs()` method to extract epochs from the raw data expects a consistent duration, we initially set `tmin = 0` and `tmax = 4`. Subsequently, we use the known duration of each condition (see Table 2) to find the end points to properly epoch the data for the technical validation steps. We retain just the data relevant to perception and imagination, and keep only the additional data related to prior visual or auditory noise/mask, for the average event related potential analyses (see subsection Average Event Related Potentials).

### Data Records

The full dataset<sup>39</sup> can be accessed at the OpenNeuro repository (<https://openneuro.org/datasets/ds004306/snapshot>). The file structure and naming follows *Brain Imaging Data Structure* (BIDS) format (<https://bids-specification.readthedocs.io/en/stable/>). See Fig. 7. The participant with visual and hearing impairments is noted in the repository.

**Raw data.** The original data produced in the ANT Neuro eego software are in `.cnt` and `.evt` format. They were converted in Matlab into `.set` and `.fdt` files to be in a format usable with the MNE package. A final conversion is computed to align the event data with BIDS format, resulting in a `.tsv` file. Therefore under the directories for each participant and session, i.e. `sub-01/ses-01/eeg/`, are four files including the raw EEG data, the electrode data, the events data and a report file. The raw data are a continuous recording of one whole session. The event files have an event label for each specific stimulus used. The trial type provides information about the specific stimulus. For example, ‘Imagination\_a\_flower\_high\_5’ refers to the imagination audio condition in which a relatively high pitched voice saying the word flower is imagined and the specific voice id of this stimulus is ‘5’. An example of a visual event is ‘Perception\_image\_flower\_c’ which refers to a perception of a flower picture. The ‘c’ indicates that the picture is relatively naturalistic/complex. Additionally, the start and end of the baseline obtained prior to the experiment tasks are provided.

**Preprocessed data.** As seen in Fig. 7, the preprocessed data is formatted as `.fff` for each participant and session. Both the EEG data and the event data can be extracted from these files in MNE. The preprocessing pipeline that has been applied to the data is described in the Data Processing section.

### Technical Validation

**Average event related potentials.** Event Related Potential (ERP) plots can be used to investigate how the brain is modulated across time in response to specific stimuli. Averaging across trials shows consistent modulations. As there is high individual variance in neural anatomy and task strategy, we calculated the average ERPs for each participant and session separately. In Fig. 8, average ERPs for each of the six tasks for participant 18 from session 1 are shown. The selected electrodes for this analysis were in occipital and posterior regions. We can see that there is no consistent pattern for imagined audio. In contrast, there is a fairly consistent ERP across electrodes for the four visual conditions.

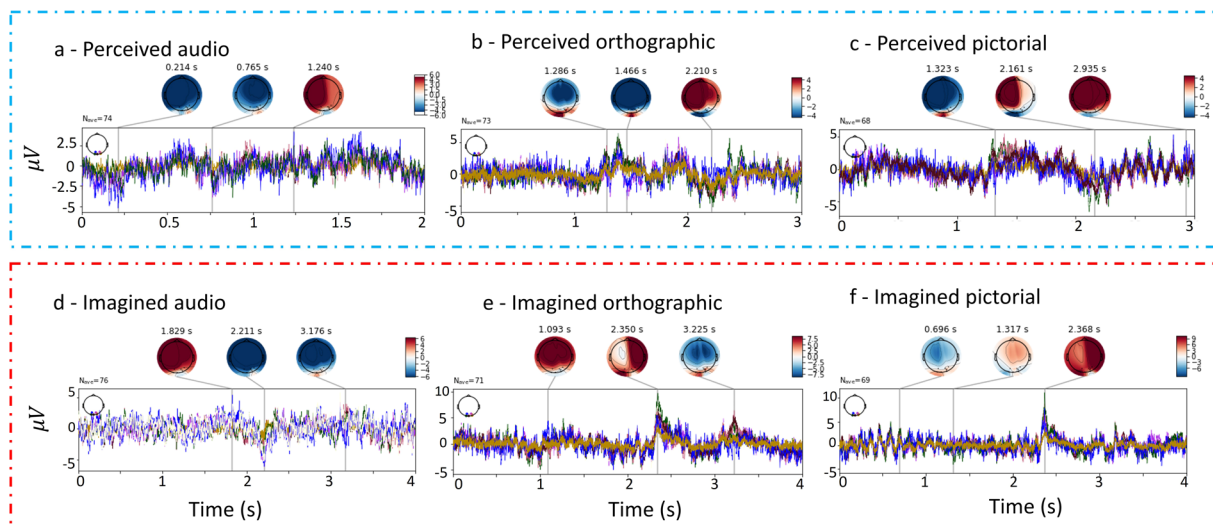
**Inter-trial coherence.** Inter-trial coherence (ITC) captures phase synchronisation or consistency across trials. A high ITC of 1 would indicate perfect coherence, whereas 0 is the lowest value and indicates no coherence. ITC is computed separately for each of the six conditions, and is shown here as the average across participants. In Fig. 9, it can be seen that there is stronger coherence for perception trials than for imagination trials. This is consistent with the expected increase of inter-trial and within participant variation in timing for generating imagined



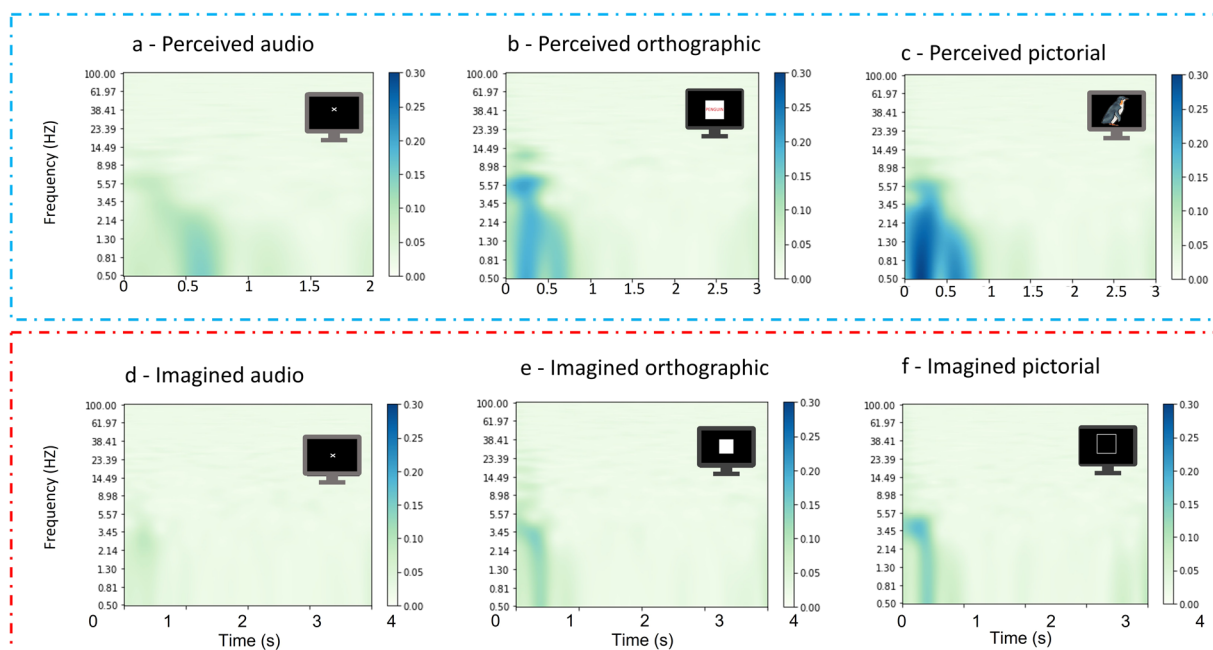
**Fig. 7** The directory structure of the data according to BIDS format. Two versions of the EEG data are provided, raw and pre-processed versions.

stimuli, whereas perceived stimuli have consistent onset and therefore higher ITC. Orthographic and pictorial perception both show strong ITC in the first 800 ms which likely relates to visual stimuli onset. Coherence is present but weaker in the same time window for imagined orthographic and pictorial tasks. Imagined audio has the least ITC, with a very weak ITC demonstrated in the first 500 ms.

**Averaged power spectral density.** We report the power spectral density (PSD) averaged over participants to represent the distribution of signal frequency components (see Fig. 10). This is computed for each of the six tasks separately. In each task there is a strong alpha peak. The plot also demonstrates that the 50 Hz power-line noise has been successfully addressed via the notch filtering described in the Filtering section.



**Fig. 8** Displaying ERP for occipital regions including the electrodes: O1, O2, O1h, O2h, I1, I2, POO9, PO8, POO9b and POO10h. This is for participant 18, session 1.



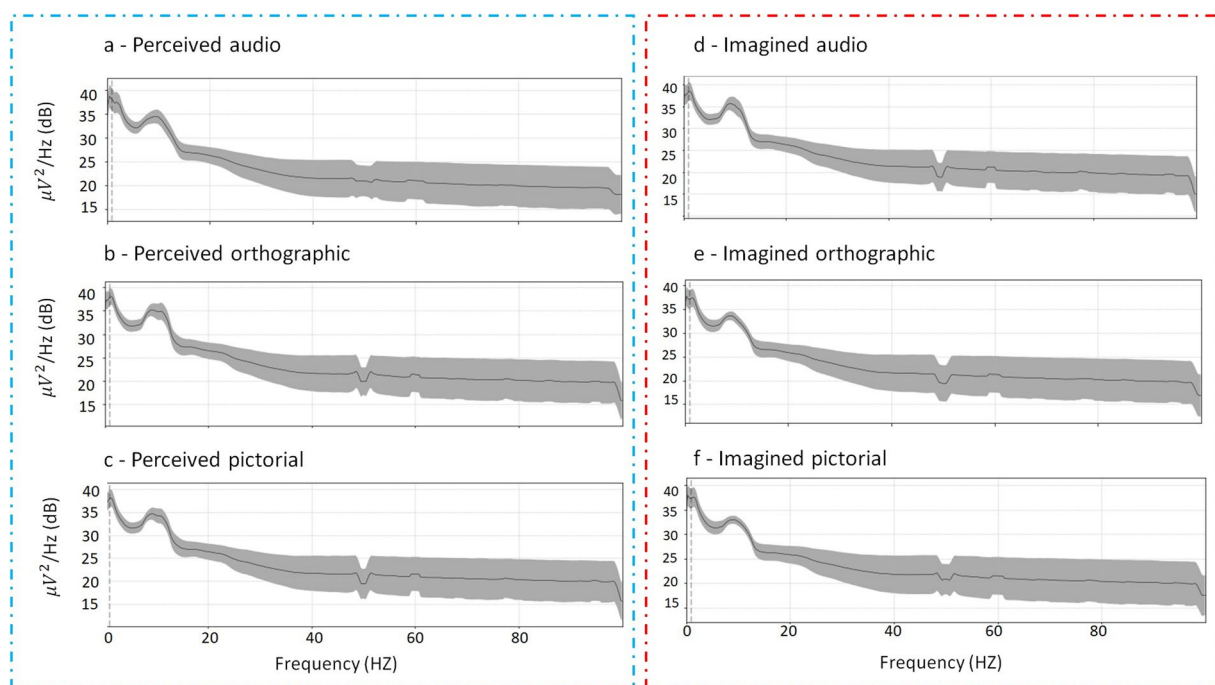
**Fig. 9** ITC for the six conditions averaged across participants. Specifically, ITC for (a) perceived audio, (b) perceived orthographic, (c) perceived pictorial, (d) imagined audio, (e) imagined orthographic and (f) imagined pictorial conditions. ITC is strongest in the perceived pictorial and orthographic conditions in the first 90 ms. ITC is weaker for imagination which is as expected due to the inter-trial variability in imagination generation and duration.

**Task classification.** To demonstrate the feasibility of this dataset<sup>39</sup> for decoding purposes and exploring neural mechanisms, we report a baseline performance on the imagination vs perception tasks for each sensory modality separately, using a logistic regression binary classification pipeline. For cognitive neuroscience, this gives insight into how distinct each task is for each modality. For BCI purposes, it can be useful to identify whether an individual is performing an imagination or perception task. To ensure consistency between the imagination and perception trials in terms of epoch length, we segmented all visual conditions into three second epochs and all the audio data into two seconds. In our analysis, we utilized a stratified cross-validation approach with five folds and conducted 50 iterations to ensure robustness of the results. The reported results (see Table 3) are averaged over the 50 iterations. Given that this was a binary classification task, chance level was set at 50%. For the visual modalities, classification accuracy is 75%. This is similar performance to that found in previous work<sup>27</sup> in which



| Task and Classifier                          | 3_3  | 8_3  | 10_1 | 11_1 | 12_1 | 12_2 | 13_1 | 14_1 | 14_2 | 15_1 | 15_2 | 16_1 | 17_1 | 18_1 | 19_1 | Avg  |
|--|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Pictorial Imagination vs. Perception (LR)    | 0.66 | 0.7  | 0.78 | 0.71 | 0.73 | 0.71 | 0.83 | 0.67 | 0.73 | 0.92 | 0.87 | 0.78 | 0.93 | 0.63 | 0.84 | 0.77 |
| Orthographic Imagination vs. Perception (LR) | 0.6  | 0.66 | 0.8  | 0.76 | 0.78 | 0.73 | 0.86 | 0.66 | 0.72 | 0.81 | 0.82 | 0.84 | 0.94 | 0.64 | 0.83 | 0.76 |
| Audio Imagination vs. Perception (LR)        | 0.48 | 0.59 | 0.59 | 0.57 | 0.62 | 0.53 | 0.63 | 0.59 | 0.55 | 0.55 | 0.63 | 0.63 | 0.72 | 0.55 | 0.71 | 0.6  |

**Table 3.** Depicting classification accuracy between imagination and perception for stratified cross validation with five folds for each participant and session, averaged over 50 iterations. LR refers to logistic regression. Here the participant number is before the underscore, and the session number after. For example, 3\_3 is participant 3 and session 3.



**Fig. 10** The average power spectral density averaged over the 124 trials and the participants for each of the six conditions (a) perceived audio, (b) perceived orthographic, (c) perceived pictorial, (d) imagined audio, (e) imagined orthographic, (f) imagined pictorial.

71% accuracy was achieved when classifying between whether a participant was imagining or observing pictures of flowers or hammers, using a SpecCSP classifier. In this current study, the average decoding performance of 60% accuracy between imagined audio and perceived audio is substantially lower. One potential explanation is that auditory perception and imagination have a higher degree of overlap than visual imagery and perception.

**Limitations and final remarks.** We present a novel high resolution EEG dataset<sup>39</sup> consisting of 124 channels. To the best of our knowledge, this is the first open source EEG dataset which captures not only semantic representation for several sensory modalities but also for both imagination and perception tasks for the same participant sample. This dataset is a promising starting point for investigating the feasibility of using semantic level representation for BCI input as well as enabling insights in cognitive neuroscience into the overlap in neural representation for semantic concepts in imagination, perception and different modalities. Still, decoding semantic representations from EEG data is difficult. To drive a representation related to semantic meaning rather than low level sensory details, we introduced high intra-class variance in this dataset. Intra-class variance results in more noise being present alongside the noise inherent from using EEG. Consequently, this is a challenging dataset for decoding, which makes it an interesting opportunity to apply deep learning techniques to extract meaningful information from noise. It is impossible to determine to what extent our participants were engaged with the experimental tasks, particularly for the imagination tasks. We included vividness metrics to indicate at minimum an individual's capacity for imagery tasks. While this metric may be relevant for the decodability of sensory information, it is less likely to correlate with semantic representation. We anticipate that decoding accuracy will vary

| Task                     | 3_3 | 8_3 | 10_1 | 11_1 | 12_1 | 12_2 | 13_1 | 14_1 | 14_2 | 15_1 | 15_2 | 16_1 | 17_1 | 18_1 | 19_1 |
|--------------------------|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Audio Perception         | 81  | 148 | 121  | 79   | 118  | 116  | 145  | 140  | 147  | 87   | 141  | 129  | 149  | 74   | 100  |
| Orthographic Perception  | 85  | 149 | 116  | 71   | 119  | 97   | 142  | 138  | 146  | 79   | 143  | 118  | 147  | 73   | 102  |
| Pictorial Perception     | 98  | 150 | 124  | 79   | 109  | 88   | 144  | 141  | 141  | 100  | 139  | 119  | 145  | 68   | 103  |
| Audio Imagination        | 85  | 150 | 114  | 75   | 117  | 111  | 146  | 142  | 149  | 84   | 133  | 122  | 150  | 76   | 104  |
| Orthographic Imagination | 76  | 150 | 118  | 64   | 118  | 98   | 147  | 140  | 148  | 78   | 137  | 118  | 144  | 69   | 103  |
| Pictorial Imagination    | 95  | 149 | 131  | 83   | 111  | 88   | 148  | 142  | 144  | 85   | 141  | 113  | 149  | 71   | 105  |

**Table 4.** Depicts the number of trials for each overall task for each participant and session. There is large variation in the number of trials completed out of 150, with a minimum of 64 and some participants completing all 150 trials. Here the participant number is before the underscore, and the session number after. For example, 3\_3 is participant 3 and session 3.

significantly between people. We hope that this dataset will create opportunities for other researchers to explore semantic decoding for BCIs as well as research questions related to neural mechanisms. By providing access not only to the raw data but also to the processed data and code for decoding, we offer a resource that can accelerate and support future research in these areas.

### Usage Notes

To facilitate the reproducibility and replicability of the study, the experiment was presented in Psychopy v.2021.2.3, a freely available software package. This ensures there are minimal barriers such as licences to prevent other researchers from using or modifying this experimental paradigm for their own studies. All code for processing and technical validation has been provided in a Jupyter Notebook tutorial style format so that following the steps for replication is as clear as possible, while also making it convenient for users to modify the code for related research questions. For example, minimal additional code is required to create classification pipelines for decoding semantics, tasks and modalities. File paths will need to be changed directly in the notebooks.

### Code availability

The Psychopy files to compile the experiment are stored on the Github repository <https://github.com/hWils/Semantics-EEG-Perception-and-Imagination>. Also on this repository are the Python processing and technical validation scripts. Users can directly use the Python code provided 1) to compute preprocessing as described in this paper, and 2) to reproduce the experimental results presented in the technical validation section.

Received: 14 February 2023; Accepted: 2 June 2023;

Published online: 15 June 2023

### References

- Nieto, N., Peterson, V., Rufiner, H. L., Kamienskowski, J. E. & Spies, R. Thinking out loud, an open-access EEG-based BCI dataset for inner speech recognition. *Sci. Data* **9**, 1–17 (2022).
- Grootswagers, T., Zhou, I., Robinson, A. K., Hebart, M. N. & Carlson, T. A. Human EEG recordings for 1,854 concepts presented in rapid serial visual presentation streams. *Sci. Data* **9**, 1–7 (2022).
- Bakhshali, M. A., Khademi, M., Ebrahimi-Moghadam, A. & Moghimi, S. EEG signal classification of imagined speech based on riemannian distance of correntropy spectral density. *Biomed. Signal Process. Control* **59**, 101899 (2020).
- Li, M., He, D., Li, C. & Qi, S. Brain-computer interface speller based on steady-state visual evoked potential: A review focusing on the stimulus paradigm and performance. *Brain Sci.* **11**, 450 (2021).
- Schlogl, A., Lee, F., Bischof, H. & Pfurtscheller, G. Characterization of four-class motor imagery EEG data for the BCI-competition 2005. *J. Neural Eng.* **2**, L14 (2005).
- Reddy, T. K., Arora, V., Gupta, V., Biswas, R. & Behera, L. EEG-based drowsiness detection with fuzzy independent phase-locking value representations using lagrangian-based deep neural networks. *IEEE Transactions on Syst. Man, Cybern. Syst.* **52**, 101–111 (2021).
- Lee, S.-H., Lee, M. & Lee, S.-W. Neural decoding of imagined speech and visual imagery as intuitive paradigms for BCI communication. *IEEE Transactions on Neural Syst. Rehabil. Eng.* **28**, 2647–2659 (2020).
- Rybák, M. & Daly, I. Neural decoding of semantic concepts: A systematic literature review. *J. Neural Eng.* (2022).
- Deniz, F., Nunez-Elizalde, A. O., Huth, A. G. & Gallant, J. L. The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *J. Neurosci.* **39**, 7722–7736 (2019).
- Arana, S., Marquand, A., Hulten, A., Hagoort, P. & Schoffelen, J.-M. Sensory modality-independent activation of the brain network for language. *J. Neurosci.* **40**, 2914–2924 (2020).
- Van de Putte, E., De Baene, W., Price, C. J. & Duyck, W. Neural overlap of l1 and l2 semantic representations across visual and auditory modalities: a decoding approach. *Neuropsychologia* **113**, 68–77 (2018).
- Simanova, I., Hagoort, P., Oostenveld, R. & Van Gerven, M. A. Modality-independent decoding of semantic information from the human brain. *Cereb. Cortex* **24**, 426–434 (2014).
- Vandenberghe, R., Price, C., Wise, R., Josephs, O. & Frackowiak, R. S. Functional anatomy of a common semantic system for words and pictures. *Nature* **383**, 254–256 (1996).
- Dijkstra, N., Bosch, S. E. & van Gerven, M. A. Shared neural mechanisms of visual perception and imagery. *Trends Cogn. Sci.* **23**, 423–434 (2019).
- Kosslyn, S. M., Ganis, G. & Thompson, W. L. Neural foundations of imagery. *Nat. Rev. Neurosci.* **2**, 635–642 (2001).
- Shen, G., Horikawa, T., Majima, K. & Kamitani, Y. Deep image reconstruction from human brain activity. *PLoS Comput. Biol.* **15**, e1006633 (2019).
- Reddy, L., Tsuchiya, N. & Serre, T. Reading the mind's eye: decoding category information during mental imagery. *NeuroImage* **50**, 818–825 (2010).
- Xie, S., Kaiser, D. & Cichy, R. M. Visual imagery and perception share neural representations in the alpha frequency band. *Curr. Biol.* **30**, 2621–2627 (2020).

19. Watanabe, H., Tanaka, H., Sakti, S. & Nakamura, S. Synchronization between overt speech envelope and EEG oscillations during imagined speech. *Neurosci. Res.* **153**, 48–55 (2020).
20. Lee, S.-H., Kravitz, D. J. & Baker, C. I. Disentangling visual imagery and perception of real-world objects. *Neuroimage* **59**, 4064–4073 (2012).
21. Breedlove, J. L., St-Yves, G., Olman, C. A. & Naselaris, T. Generative feedback explains distinct brain activity codes for seen and mental images. *Curr. Biol.* **30**, 2211–2224 (2020).
22. Gaziv, G. *et al.* Self-supervised natural image reconstruction and large-scale semantic classification from brain activity. *NeuroImage* **254**, 119121 (2022).
23. Qiao, K. *et al.* Category decoding of visual stimuli from human brain activity using a bidirectional recurrent neural network to simulate bidirectional information flows in human visual cortices. *Front. Neurosci.* **13**, 692 (2019).
24. Sheikh, U. A., Carreiras, M. & Soto, D. Decoding the meaning of unconsciously processed words using fMRI-based MVPA. *NeuroImage* **191**, 430–440 (2019).
25. Kivisaari, S. L. *et al.* Reconstructing meaning from bits of information. *Nat. Commun.* **10**, 1–11 (2019).
26. Murphy, B. & Poesio, M. Detecting semantic category in simultaneous EEG/MEG recordings. In *Proceedings of the NAACL HLT 2010 First Workshop on Computational Neuro Linguistics*, 36–44 (2010).
27. Kosmyna, N., Tarpin-Bernard, F. & Rivet, B. Operationalization of conceptual imagery for BCIs. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, 2726–2730 (IEEE, 2015).
28. Simanova, I., Van Gerven, M., Oostenveld, R. & Hagoort, P. Identifying object categories from event-related EEG: toward decoding of conceptual representations. *PLoS One* **5**, e14465 (2010).
29. Fulford, J. *et al.* The neural correlates of visual imagery vividness—an fMRI study and literature review. *Cortex* **105**, 26–40 (2018).
30. Zabicki, A. *et al.* Subjective vividness of motor imagery has a neural signature in human premotor and parietal cortex. *NeuroImage* **197**, 273–283 (2019).
31. Halpern, A. R. Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music. Mind Brain* **25**, 37 (2015).
32. Marks, D. F. Vividness of visual imagery questionnaire. *J. Mental Imag.* (1973).
33. Peirce, J. *et al.* Psychopy2: Experiments in behavior made easy. *Behav. Res. Methods* **51**, 195–203 (2019).
34. Oostenveld, R. & Praamstra, P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* **112**, 713–719 (2001).
35. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781* (2013).
36. Harel, A., Gilaie-Dotan, S., Malach, R. & Bentin, S. Top-down engagement modulates the neural expressions of visual expertise. *Cereb. Cortex* **20**, 2304–2318 (2010).
37. Chen, L., Cichy, R. M. & Kaiser, D. Semantic scene-object consistency modulates n300/400 EEG components, but does not automatically facilitate object representations. *Cereb. Cortex* **32**, 3553–3567 (2022).
38. Klug, M. & Gramann, K. Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments. *Eur. J. Neurosci.* **54**, 8406–8420 (2021).
39. Wilson, H., Golbabaee, M., Proulx, M. & O'Neill, E. *EEG semantic imagination and perception dataset*, <https://doi.org/10.18112/openneuro.ds004306.v1.0.1> (2023).

## Acknowledgements

HLW's research is supported by EPSRC grant EP/S515279/1. MG's research is supported by the EPSRC grant EP/X001091/1. MJP's research is supported by the AHRC grant AH/T004673/1. EON's research is supported by CAMERA 2.0, the EPSRC Centre for the Analysis of Motion, Entertainment Research and Applications (EPSRC grant EP/T022523/1). We are very grateful to Aneekha Bal, Becky Dakin and Jinha Yoon for their assistance with data collection. MJP is also a Research Scientist at Meta Reality Labs Research.

## Author contributions

HLW designed and conducted the experiments, completed pre-processing, computed analyses and designed classification pipelines. SC. was involved in the pre-processing and design of classification pipelines. All authors were involved in writing and revising the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to H.W. or E.O.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023