Source Reconstruction of the ERP Uncertainty Effects reveals common Neural Mechanisms for different Stimulus Categories

Lukas Hecker (lukas_hecker@web.de)
Department of Psychiatry and Psychotherapy, University of Freiburg Medical Center, Freiburg
Ellen Joos
INSERM U1114, Cognitive Neuropsychology and Pathophysiology of Schizophrenia, Strasbourg
Bernd Feige
Department of Psychiatry and Psychotherapy, University of Freiburg Medical Center, Freiburg
Simon Maier
Department of Psychiatry and Psychotherapy, University of Freiburg Medical Center, Freiburg
Ludger Tebartz van Elst
Department of Psychiatry and Psychotherapy, University of Freiburg Medical Center, Freiburg
Jürgen Kornmeier
Perception and Cognition Lab, Institute for Frontier Areas of Psychology and Mental Health, Freiburg

Article

Keywords:

Posted Date: January 12th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2405042/v1

Additional Declarations: No competing interests reported.

Source Reconstruction of the ERP Uncertainty Effects reveals common Neural Mechanisms for different Stimulus Categories

Hecker, Lukas[1,2,3,4,5]; Joos, Ellen[6]; Feige, Bernd[5]; Maier, Simon[1]; Tebartz van Elst, Ludger[1]; Kornmeier, Jürgen[1,3,4,5]

[1] Department of Psychiatry and Psychotherapy, University of Freiburg Medical Center, Freiburg, Germany.
[2] Department of Psychosomatic Medicine and Psychotherapy, University of Freiburg Medical Center, Freiburg, Germany.
[3] Perception and Cognition Lab, Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany.
[4] Faculty of Biology, University of Freiburg, Freiburg, Germany.
[5] Medical Center - University of Freiburg, Freiburg, Germany.
[6] INSERM U1114, Cognitive Neuropsychology and Pathophysiology of Schizophrenia, Strasbourg, France

*Corresponding author: lukas.hecker@uniklinik-freiburg.de

- **Funding:**
- **Conflict of interest:** The authors declare that there is no conflict of interest regarding the publication of this article.

1 Abstract

During the observation of an ambiguous figure, our perception becomes unstable and alternates repeatedly between mutual exclusive interpretations. Tiny changes of the stimulus features can disambiguate the figure and stabilize percepts. Recent EEG studies found much smaller amplitudes of two event-related potentials (ERPs), an anterior P200, 200 ms after stimulus onset and a posterior P400, 200 ms later, when participants observed an ambiguous stimulus compared to disambiguated stimulus variants. Interestingly, this pattern of results was found across stimuli differing in ambiguity (geometry, motion and gestalt) and in visibility. We postulate a meta-perceptual / meta-cognitive evaluation instance that rates the reliability of perceptual constructs at a high processing level generalized beyond sensory details. We further postulate that the above described ERP effects reflect the outcome of this evaluation process.

According to these hypotheses, the distributions of these ERP effects on the scalp across three different stimulus categories, should originate from the same neural structures in the sensor and source space. This was tested by calculating EEG inverse solutions using a novel artificial neural network - based approach. We found very similar sources across stimulus categories, both on the level of individual participants and on the group level. Regions involved in the earlier processing steps (P200) encompass lateral occipital cortex, inferior parietal cortex and medial cingulate cortex. Later processes (P400) originated mostly from the right inferior temporal cortex. Our findings were consistent with comparable studies using functional magnetic resonance imaging and EEG source imaging. In summary, we found highly coherent neural sources of the ERP Uncertainty Effects. The underlying processes may be related to a common uncertainty resolution at a higher processing level beyond sensory details.

2 Introduction

The information available to our senses is incomplete and to varying degrees ambiguous. Our perceptual system needs to disambiguate and interpret this limited information in order to produce stable and reliable perceptual interpretations as fast as possible. Ambiguous figures, like the Necker Cube [1] or Boring's Old/Young Woman [2], are extreme cases, where two (or more) valid interpretations are about equally probable for one and the same visual information, leading to only transiently stable percepts with short transitions (reversals) from one interpretation to another (multistability). Ambiguous figures were used extensively in the past decades to study our perceptual system in various contexts. The majority of the work was focused on the mechanisms underlying perceptual reversals (e.g., 3, 4, 5, 6). A different focus on the topic was set by a number of more recent EEG studies [7, 8]. The authors were interested in the differences between stable neural representations of disambiguated and unstable representations of ambiguous stimuli They found significantly smaller amplitudes of two event-related potentials (ERPs), a more anterior P200 and a more posterior P400.

Interestingly, this pattern of results generalized across different categories of ambiguous stimuli, like the Necker cube, von Schiller's Stroboscopic Alternative Motion ("SAM", also known as the motion quartet; 9) and Boring's old/young woman. Besides ambiguous figures, very similar effects are also found for stimuli of varying visibility, e.g., concerning the emotional expression of faces [8, 10]. Since these ERP effects can be observed for both ambiguous and low-visibility stimuli they were termed "ERP Uncertainty Effects" (see 8 for a discussion on this).

This generalization of the ERP Uncertainty Effects indicates common neural mechanism of ambiguity or uncertainty processing at a meta-cognitive / meta-perceptual level of processing beyond low-level feature extraction. In the present work we analyzed the neural sources of the ERP Uncertainty Effects evoked by ambiguous/ low visibility stimuli and

disambiguated/ high visibility stimulus variants from three very different categories. We studied whether the apparently similar scalp distributions of ERPs across stimulus categories are based on the same underlying neural sources.

The present work aims to identify the neural generators underlying these meta-perceptual ERP Uncertainty Effects for a set of three different stimuli: Necker Lattices, smileys and abstract figures. The data was previously analyzed and published by [11]. A classical ERP analysis revealed highly similar Uncertainty Effects for the three stimulus types. In the present study, we investigate whether highly similar amplitude effects combined with highly similar scalp distributions of the related ERP signatures are based on the same neural sources, confirming our hypothesis. In order to test this, we applied an advanced EEG source analysis method based on artificial neural networks (ANNs, 12, 13). ANNs have shown better source localization accuracy compared to classical, linear approaches to the inverse problems such as exact low-resolution tomography (eLORETA, 14). One important advantage has been the ability for ANNs to not only estimate the location but also the spatial extent of neural sources. However, since the ANNs are a relatively new re-emerging method in the field, we have decided to also calculate the established linear inverse solution local autoregressive average (LAURA, 15) as a complement.

The majority of existing studies on neural sources of ambiguity processing are based on functional magnetic resonance imaging (fMRI) and mainly focused on the reversal process (for an extensive review, see 3). Only few studies have contrasted correlates of neural representations during periods of stable percepts and compared these between ambiguous stimuli and disambiguated variants thereof (e.g., 16). None of those studies had a focus on the relation between ambiguity and visibility with a specific focus on uncertainty processing and particularly on the Uncertainty Effects reported by our lab [10, 8, 17, 18]. We therefore have no a priori spatial regions of interest (ROI) for our analysis and thus followed a mainly explorative approach. However, we expect that: (1) there will be an overlap of

the ambiguity/uncertainty effects of all three stimuli used in the present study due to the similarity in scalp distributions of the ERP Uncertainty Effects. (2) This overlap will be larger between the smileys and abstract figures because of the larger stimulus similarity.

3 Methods

3.1 Participants

Twenty participants (9 male) between 19 and 34 years (median age: 24 years) with normal or corrected-to-normal visual acuity, as measured with the FrACT [19] took part in the present study. All participants gave their informed written consent. The study was approved by the ethics committee of the University of Freiburg and in accordance with the ethical standards laid down in the Declaration of Helsinki (SRC, [20]).

3.2 Stimuli

The stimulus set consisted of three ambiguous figures and their unambiguous variants, namely Necker Lattices, Smileys and Abstract Figures (see Fig. 1). The Necker Lattices are construed of 3×3 stacked up Necker Cubes [1, 21]. Disambiguated lattice variants (Fig. 1, top left), representing the two perceptual alternatives of the ambiguous lattice, were created by adding depth cues like shading, central projection and aerial perspective (OpenGL lighting model; 22). All lattices extended over a visual angle (VA) of 7.5×7.5.

Interestingly, [8] found recently highly similar ERP Uncertainty Effects for ambiguity/uncertainty in the emotional expression of smiley faces, but also for uncertainty concerning low-level visibility of curvature elements of an abstract figure. The Smileys depict a minimalistic facial expression of a face with happy or a sad emotional expression (Fig 1, center column). This facial expression was varied by simply changing the radius and orientation of the mouth curvature. Uncertainty with respect to emotion can be evoked

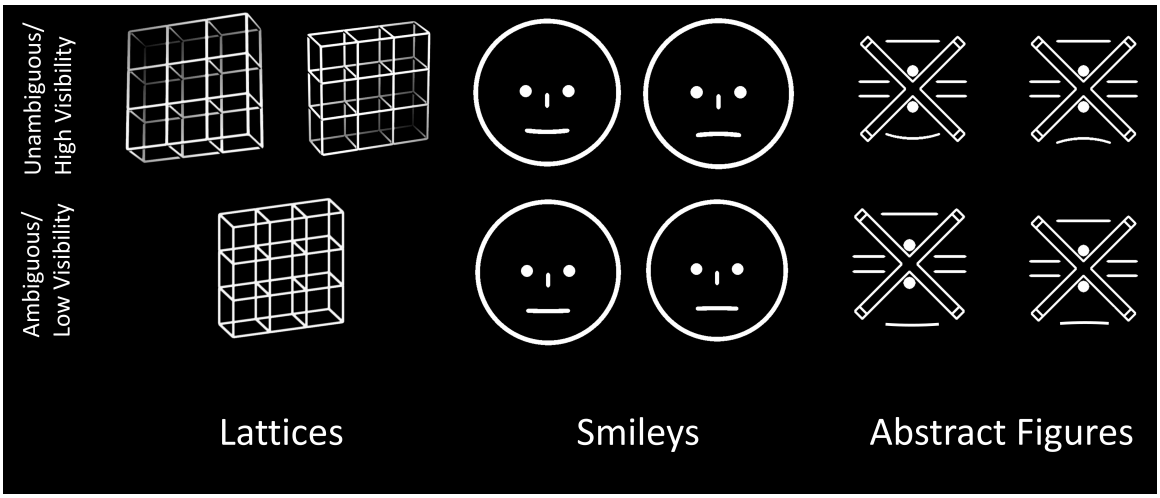


Figure 1: **Stimulus Set** Top row: Unambiguous Versions. Bottom row: Ambiguous version. Ambiguity in Necker Lattices exists with regard to stimulus orientation. Smileys are ambiguous regarding their emotional expression. Abstract Figures show ambiguity regarding low-level features of line curvature on the bottom.

by increasing the radius of the mouth curvature, which flattens it. All Smileys were sized to extend over 4×4 VA. Abstract Figures contained the same mouth curvature lines as the smileys, which were however embedded in abstract structures, containing also mirror symmetry and the same total sum of white area as the smileys but being otherwise very dissimilar to faces (Fig. 1, right column).

3.3 Procedure

In two separate experimental sessions, participants completed in total six different experimental conditions including ambiguous and unambiguous variants of Necker lattices. The two sessions took place on two different days that were maximally one week apart from each other. Abstract figures were always shown in the first session and smileys in the second session to avoid an association between the abstract figures and a facial stimulus. Each condition was split up into three blocks of 7 minutes (lattices) or 6 minutes (smileys and abstract figures) each. The order of experimental blocks and the presentation order within

the blocks was randomized. Participants sat in front of a cathode ray tube with 1.14 m distance between the eye and the screen. Luminance of all stimuli was kept at $40cd/m^2$. Stimuli were shown discontinuously (see 23) for 1 second, followed by a dark screen that persisted for 0.4 seconds. Participants were asked to compare their current percept with the percept during the previous trial by button press. They indicated stability if the percept remained the same and reversal if it changed since the last trial. For lattices the two possible perceptions were front-side pointing upwards or downwards (Fig. 1, top left). Smileys could be perceived as either happy or sad. For the abstract figures the bent line could be perceived to point either upwards or downwards. Participants were instructed to not focus on the line curvature of ambiguous smileys but instead try to respond to the subtle changes in the perceived emotion (wholistic perception of face and emotion).

3.4 EEG Recording

EEG was recorded using 32 active electrodes according to the international 10-20 system (ActiCap) and an ActiChamp amplifier (Brain Product). Data was sampled at 1000 Hz and online band-pass filtered from 0.01 to 120 Hz.

3.5 EEG Preprocessing

All preprocessing of the EEG data was accomplished using mne-python (v1.0.2, 24). All code used for the analyses described henceforth is publicly available at `https://github.com/LukeTheHecker/ambiguity_sources/`.

Raw data was referenced to common average and filtered using a FIR bandpass from 0.01 - 45 Hz. In some cases, line noise remained even after filtering. Therefore, we removed line noise at 50 Hz and all harmonics up until 500 Hz using a notch filter using a finite impulse response (FIR) filter and a 1 Hz width of the transition band.

Eye artifacts were removed using independent component analysis, specifically the PI-

CARD method [25, 26]. Next, bad channels were detected and interpolated using random sampling consensus algorithm as implements by the autoreject package for python [27, 28].

This procedure yielded clean EEG data which was then epoched into trials of the different conditions. Finally, we rejected all trials that contained amplitudes exceeding $150\mu V$.

3.6 Forward Model

Standard channel positions were assumed for all $q = 32$ electrodes according to 10-20 System since no true channel positions were available from the data set. We used the "fsaverage" [29] template T1 image as provided by the Freesurfer image analysis suite (https://surfer.nmr.mgh.harvard.edu/). The forward model was computed using the boundary element method (BEM, 30) as provided by mne-python. Each shell (brain, skull & scalp tissue) was composed of 5120 vertices. Conductivity was set to $0.3S/m$ for brain and scalp tissue, and $0.06S/m$ for the skull.

The source space was created using $p = 1284$ dipoles with icosahedral spacing. A fixed orientation of dipoles orthogonal to the surface of the cortical sheet was assumed due to the reasonable physiological assumption and in order to reduce computational complexity.

3.7 Source Analysis

3.7.1 Artificial Neural Network (ANN)

Neural sources underlying the ERPs were calculated using a fully-connected ANN, which has shown to yield accurate inverse solutions with respect to source location and estimation of source extent [31, 13]. We used simulated neural sources and scalp EEG data to train the fully-connected (FC) ANN model. The esinet package (v.0.2.4, https://github.com/LukeTheHecker/esinet) for python was used to simulate neural sources and corresponding EEG data and to build and train the FC model.

We simulated 20,000 samples of source and EEG activity. Each sample contained brain

activity of 1 - 15 randomly located and sized sources spanning 50 time points. The activity of each source was determined using noise that followed the frequency spectrum $f^{-\beta}$, where β is the slope of the $\frac{1}{f}$ spectrum. This slope was calculated for the ERP of each participant's EEG data and found in the range of 1.3 to 1.8. A value within this range was randomly selected within the range for each simulated sample. Signal-to-noise ratio (SNR) was also estimated from the ERPs of each participant and condition. This was achieved by first calculating the global field power (GFP) for each participant and condition. Then we calculated the ratio between the maximum of the GFP and the averaged GFP during a baseline interval from 200 ms pre-stimulus until stimulus onset. We found that the SNR of our participants ranged from 4.2 - 19.9, which was again used as the parameter range for the simulation. For further details on the simulation procedure please refer to [13].

The ANN was built and trained using tensorflow 2.5.0 [32] and keras 2.5.0 [33]. The FC was composed of an input layer with 32 neurons, corresponding to the $q = 32$ EEG electrodes. Two hidden layers were followed with 300 and 500 neurons, respectively. Tanh activation functions were applied to each hidden layer and dropout was applied during training at a rate of 20 %. The output layer consisted of 1284 neurons, corresponding to the $p = 1284$ dipoles in our source model (i.e., positions in the brain). The model contained a total number of 803,684 parameters. The model was trained on the simulated EEG and source data for 75 epochs using a batch size of 8. We used the adaptive momentum estimation (ADAM, 34) optimizer with default parameters. As a loss function we used the cosine distance, which calculates the error independent of scaling (similar to a correlation coefficient).

3.7.2 Local Auto-Regressive Average (LAURA)

We decided to complement the inverse solutions yielded from the LSTM model with a those from a simpler linear inverse solution and see if the results of the two approaches are approx-

imately aligned. We therefore decided for Local Auto-Regressive Average (LAURA, 15), a linear approach to solve the M/EEG inverse problem using biophysical constraints. The constraints include the assumption on electromagnetic fields in biological media where magnitude of current decreases proportionally with the squared distance. We have decided for LAURA since it is capable to localize multiple neural sources and showed good localization accuracy in our tests. Furthermore, it has been used to find the neural generators of Necker cube reversals [35] already, which allows for a more direct comparison. The LAURA inverse solutions were calculated using the *invert package*[https://github.com/LukeTheHecker/invert]. The optimal regularization parameter was found using generalized cross validation as described by [36].

3.8 Similarity Analysis

The core of this study is to determine whether similar ERP Uncertainty Effects identified with our three stimulus categories are based on similar neural sources. Similarity will be measured on two levels: On the level of individual participants and on the level of the grand-average.

The individual similarity was calculated as follows: First, the sources underlying the ERPs to all stimuli (lattice, smiley and abstract figures) and conditions (unambiguous/high visibility and ambiguous/ low visibility) were calculated using the LSTM model. Then we calculated the difference source for each participant and stimulus (*unambiguous/high visibility* − *ambiguous/low visibility*). Finally, we calculated the Pearson correlation coefficient between the source vectors of each pair of stimuli at each time point. This results in a time-resolved source similarity measure per subject and for the three stimulus-pairs. Using this approach, we lose the information of *where* the effects are located in the brain but gain a clearer picture of *when* the processing of ambiguity/uncertainty is similar among two different stimulus categories.

The grand-average analysis allows us to localize sources on the group-level and requires the calculation of the sources of the uncertainty effects per participant as a first step. For each stimulus type (lattice, smiley & abstract figure) we calculated the difference between the unambiguous/high visibility and ambiguous/low visibility condition by subtraction. Next, the resulting difference was normalized per subject in order to reduce inter-individual differences in amplitudes. Normalization was calculated by dividing the difference source matrix by its own mean absolute deviation.

A grand average was then calculated by averaging the normalized source differences across subjects for each stimulus, individually.

3.9 Source Statistics

Source activation maxima were identified in the grand-average at the two components of interest, P200 (150-300 ms) and P400 (300-600 ms). First, the values below the 10th percentile were set to zero in order remove sources that contribute little to the overall pattern. Next, we identified local maxima in the grand-average. We defined a maximum as a vertex whose 5 closest neighbors had lower values. Finally, the resulting maxima were filtered when there was a larger maximum within a radius of 30 mm. This filtering removed *local maxima*.

Statistic was then calculated based on the individual's values at the location of the aforementioned maxima. Statistical significance of voxel activations were assessed using a two-sided one-sample Wilcoxon test against the null hypothesis H_0 of no significant deviation from 0. The anatomical regions of the maxima were described using the Multi-Modal Parcellation of the *Human Connectome Project* (HCP-MMP1, 37) and using classic Brodmann areas [38]. Coordinates of the maxima were reported in MNI-space.

4 Results

4.1 Event-Related Potentials (ERPs)

Fig. 2 shows the ERPs in response to each stimulus, separately. Responses are divided into unambiguous/ high visibility, ambiguous/ low visibility and their difference ERP (dERP). The ERPs largely resemble those found in the previous analysis of the present data (cf. 8, Fig. 4), apart from the different EEG reference. We find that each stimulus evokes a P200 component, which is larger in the unambiguous/ high visibility stimulus.

Similarly, all stimuli evoked a P400 component which extends from 300 to 600 ms after stimulus onset. The topographies of the difference P200 (dP200) components are similar between smileys and abstract figures, whereas the lattice's dP200 shows a dipolar pattern peaking at frontal and occipito-temporal electrodes distinct to that of the other two stimuli. On the contrary, the dP400 is largely similar across all different stimuli.

We have quantified the similarity of difference-based topographies between each stimulus pair by calculating subject-wise and time-point-wise Pearson correlation coefficients (see Similarity Analysis) as depicted in Fig. 3a. It shows that the individual dERP topographies are highly correlated after stimulus onset between all three pairs of stimuli. This correlation is especially strong starting from 150ms after stimulus onset and declines slowly after a peak between 300 and 600 ms.

4.2 Source Analysis

While the similarity between individual dERPs of the different stimulus categories was considerably high, it is unclear whether these similarities originate from the same regions in the brain. We have conducted a source analysis of the individual ERPs using two different approaches in order to get a more reliable estimate: Using the FC network and using the linear inverse solution LAURA. The former will henceforth be presented in the main results

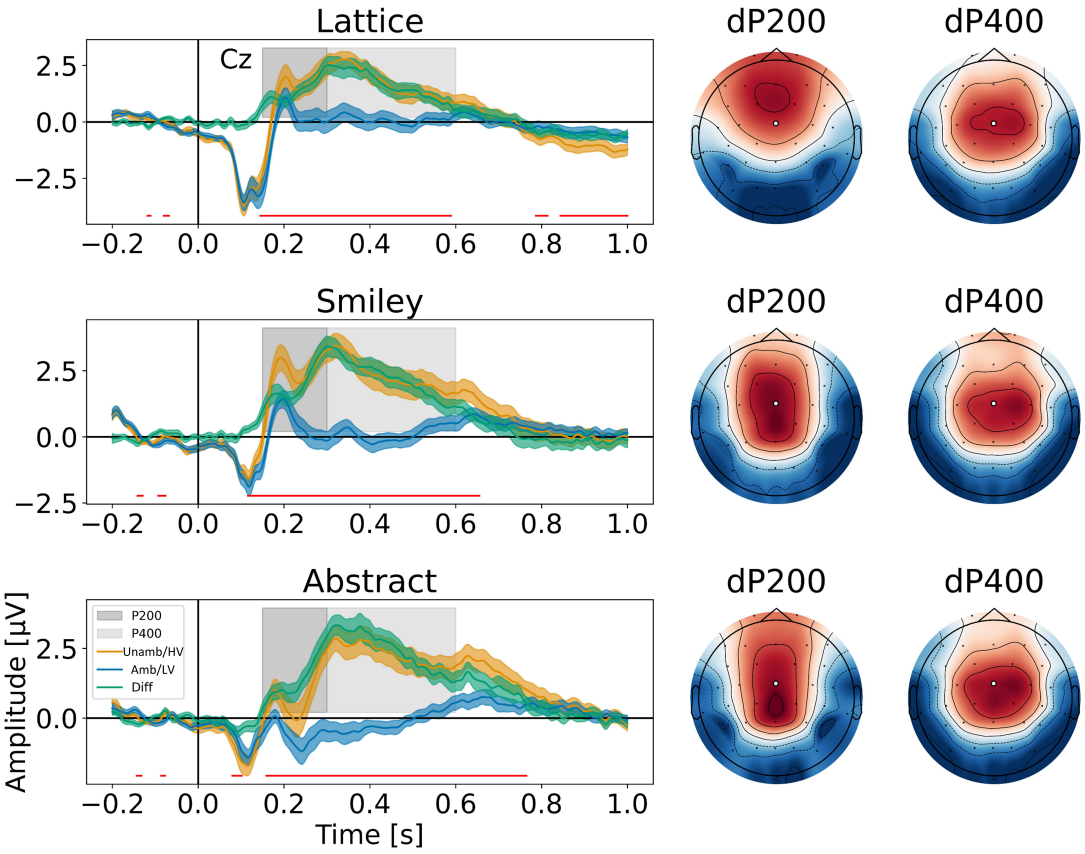
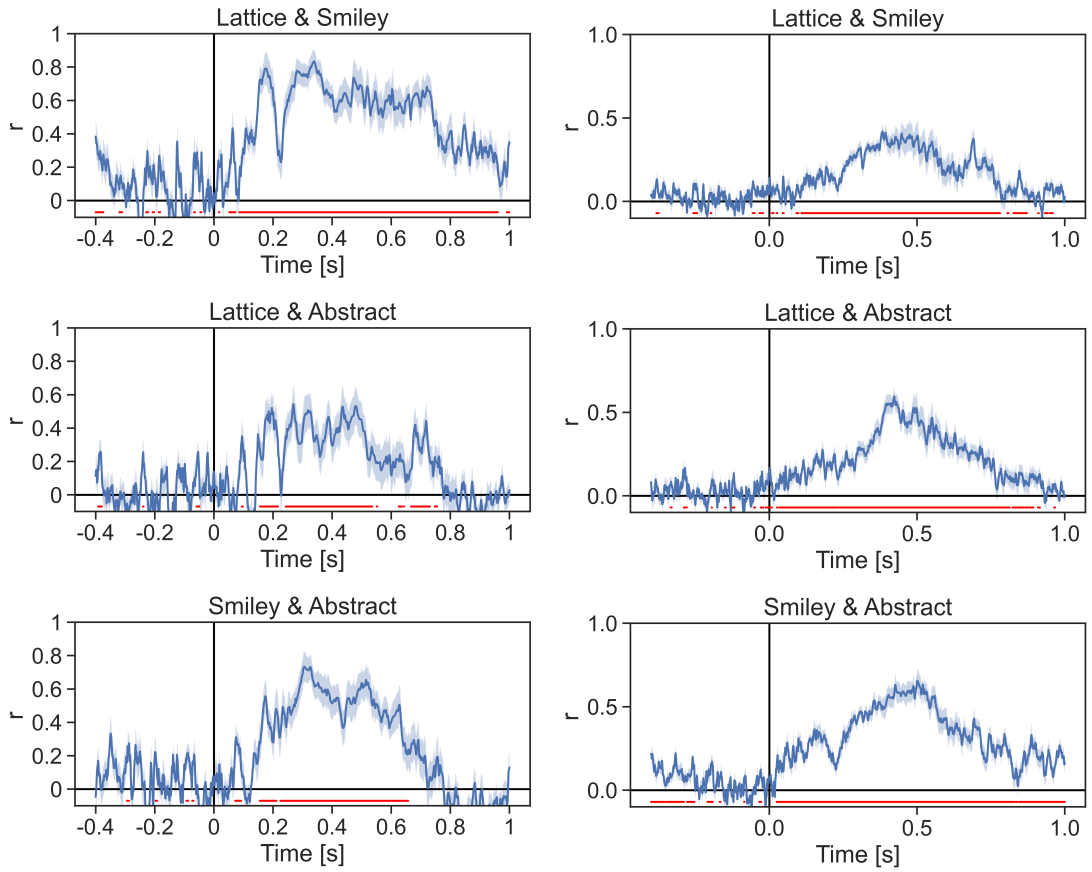


Figure 2: **Event-Related Potentials (ERPs)** Left column shows the ERPs unambiguous/ high visibility (orange) and ambiguous/ low visibility (blue) stimuli and the difference unambiguous/ high visibility minus ambiguous/ low visibility (green) at the vertex electrode Cz. The topographical plots on the right column show the voltage distribution of the difference ERP (dERP) at our temporal regions of interest (tROI) with the vertex electrode highlighted in white. The tROI of the P200 spans from 150 ms to 300 ms after stimulus onset, whereas the P400 spans from 300 ms to 600 ms post-stimulus (see area shaded in grey). HV: High-visibility stimulus. LV: Low-visibility stimulus.



(a) **Electrode-Space** (b) **Source-Space (FC)**

Figure 3: **Time-Resolved Individual Similarity** Grand-average of the pair-wise Pearson correlation between differential neural responses to unambiguous/high visibility vs. ambiguous/low visibility stimuli. (a) Topographic similarity (electrode-space analysis). (b) Source similarity (source-space analysis). Top row: Lattice difference-ERPs (unambiguous - ambiguous) compared to smiley dERPs (high visibility - low visibility). Middle row: Lattice dERPs (unambiguous - ambiguous) compared to abstract figure dERPs (high visibility - low visibility). Bottom row: Smiley dERPs (high visibility - low visibility) compared to abstract figure dERPs (high visibility - low visibility). Shadings indicate standard errors. Red line indicates significant correlations per time point tested across participants.

section. See Appendix A for equivalent analyses using LAURA, showing a comparative pattern.

The source-space similarity (Fig. 3b) is highly significant, yet overall lower compared to the electrode-space similarity (Fig. 3a). In source-space, uncertainty effects in lattice and smiley stimuli show a significant, albeit lower similarity overall with correlation coefficients $r <= 0.38$ peaking at 500 ms after stimulus onset. In stark contrast, the source similarity between smileys and abstract figures is considerably higher in the source space, indicating that uncertainty resolution is a highly similar process with these two stimulus categories. The correlation is significant 80 ms after stimulus onset until stimulus offset, peaking at 500 ms with at $r = 0.56$. The similarity between lattice stimuli and abstract figures sits between the aforementioned contrasts, peaking at 400 ms with $r = 0.50$. A highly similar pattern was observed using the LAURA inverse solution (see Appendix A).

As described earlier, we also calculated grand-averages of the source-space uncertainty effects for the dP200 and dP400, respectively. The grand-average source estimates of the dP200 are visualized in Fig. 4 and the list of maxima is found in Table 1.

We identified multiple clusters of ambiguity-/uncertainty-related neural sources that are shared between all stimulus categories during the P200 time window (see Tab. 1). The largest clusters encompass the bilateral lateral occipital cortex (LOC) and the right inferior/middle temporal cortex (IT/MT) (see *Intersection* in Fig. 4). A scattered cluster of shared generators was found in the left postcentral gyrus and the middle cingulate cortex (MCC). Using LAURA, we can see a highly similar pattern of intersecting brain regions of the dP200 uncertainty effects (Appendix A), including bi-lateral LOC, right IT and MCC.

The sources of the dP400 show broader activations, which may in part be due to stronger effects in that time window (cf. Fig. 2 & Tab. 2) leading to more significant differences in source space. We find a large overlapping cluster throughout the right lateral and medial temporal cortex. Each stimulus, however, showed that the peak of this broad activity is

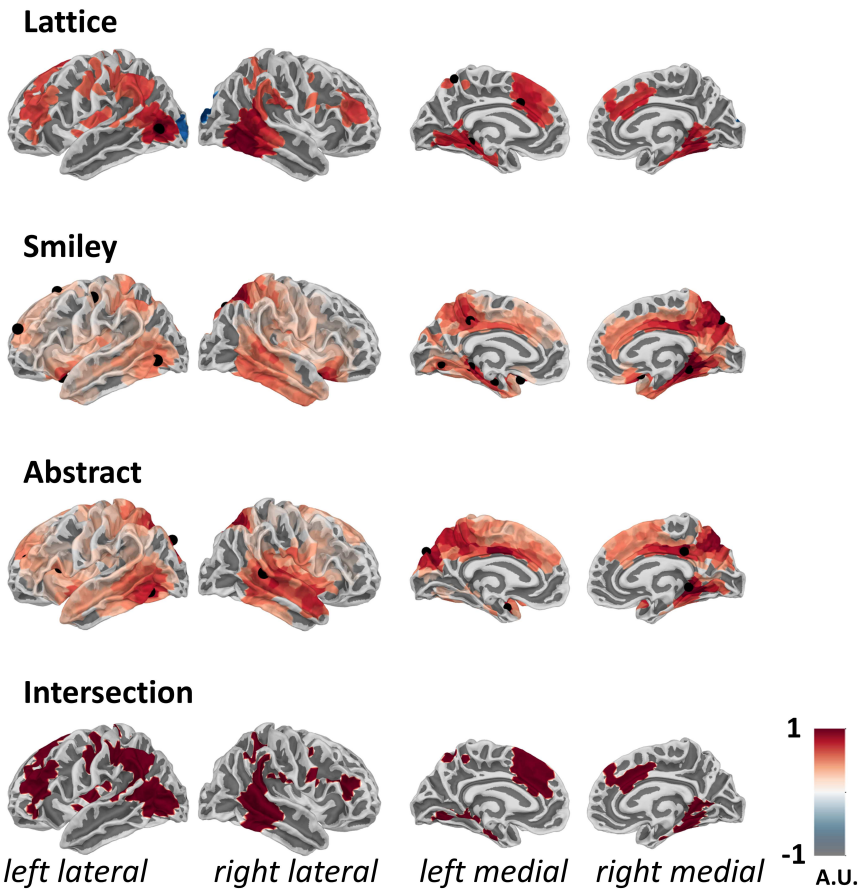


Figure 4: **FC-Sources of the P200** Grand average of the FC source differences unambiguous/high visibility - ambiguous/low visibility averaged across the P200 time range (150 - 300 ms post-stimulus) are shown. Units are arbitrary due to normalization of the individual differences. Only vertices that showed a significant difference at $\alpha < 0.05$ are shown. The intersection shows a binary mask of the vertices that were significantly active in all three stimuli. Sources in the left occipito-temporal and right lateral temporal cortex show larger activity for the unambiguous/ high-visibility stimuli throughout all categories.

Stimulus	HCPMMP1	BA	Hem	X	Y	Z	T	p
Lattice								
	MT+ Complex and Neighboring Visual Areas	BA 19	L	-39.8	-68.1	-0.0	4.21	0.0005
	Anterior Cingulate and Medial Prefrontal Cortex	BA 24	L	-11.8	14.2	34.3	2.96	0.0080
	Lateral Temporal Cortex	BA 37	R	42.0	-50.5	-12.0	2.89	0.0094
	Dorsal Stream Visual Cortex	BA 18	R	23.9	-80.2	18.5	-2.73	0.0133
	Superior Parietal Cortex	BA 7	L	-9.2	-54.8	59.6	2.32	0.0318
Smiley								
	Posterior Cingulate Cortex	BA 31	L	-10.7	-38.4	41.8	5.13	0.0001
	Orbital and Polar Frontal Cortex	BA 11	L	-16.9	14.5	-20.6	4.88	0.0001
	Early Visual Cortex	BA 18	L	-14.7	-70.3	-3.9	4.73	0.0001
	Somatosensory and Motor Cortex	BA 4	L	-40.8	-11.8	48.1	3.95	0.0009
	Medial Temporal Cortex	BA 36	L	-34.5	-13.8	-25.1	3.88	0.0010
	Posterior Cingulate Cortex	BA 7	R	17.8	-75.4	43.4	3.53	0.0022
	MT+ Complex and Neighboring Visual Areas	BA 37	L	-46.8	-64.8	-5.1	3.44	0.0028
	DorsoLateral Prefrontal Cortex	BA 6	L	-16.6	19.7	57.2	3.35	0.0034
	Superior Parietal Cortex	BA 7	L	-45.7	-41.5	40.8	3.26	0.0041
	DorsoLateral Prefrontal Cortex	BA 9	L	-22.6	53.7	22.3	2.86	0.0101
Abstract								
	MT+ Complex and Neighboring Visual Areas	BA 37	L	-40.2	-60.1	-10.6	4.56	0.0002
	Temporo-Parieto-Occipital Junction	BA 22	R	46.0	-36.9	5.2	4.34	0.0004
	Posterior Cingulate Cortex	BA 23	R	6.3	-34.8	34.5	3.95	0.0009
	Dorsal Stream Visual Cortex	BA 18	L	-16.5	-86.6	36.1	3.19	0.0048
	Medial Temporal Cortex	BA 35	R	15.0	-40.3	-2.3	2.70	0.0142
	Dorsal Stream Visual Cortex	BA 19	R	22.1	-71.1	40.6	2.60	0.0176
	DorsoLateral Prefrontal Cortex	BA 10	L	-26.3	45.0	16.4	2.54	0.0201
	Lateral Temporal Cortex	BA 38	L	-34.8	0.4	-25.0	2.40	0.0268

Table 1: **P200 Maxima** The source maxima of at the P200 time window 150 ms - 300 ms after stimulus onset. HCP-MMP1: The Multi-Modal Parcellation atlas of the Human Connectome Project. BA: Brodmann Areas, X,Y,Z: MNI coordinates, T: T-value of the Wilcoxon test. p: corresponding uncorrected p-value of the Wilcoxon test.

Stimulus	HCPMMP1	BA	Hem	X	Y	Z	T	p
Lattice								
	Superior Parietal Cortex	BA 2	L	-39.5	-34.8	35.4	3.80	0.0012
	Somatosensory and Motor Cortex	BA 3	R	40.7	-21.3	46.9	2.97	0.0078
	Paracentral Lobular and Mid Cingulate Cortex	BA 6	R	12.3	-0.1	46.8	2.88	0.0096
	DorsoLateral Prefrontal Cortex	BA 9	L	-15.5	45.7	38.6	2.59	0.0181
	Paracentral Lobular and Mid Cingulate Cortex	BA 5	R	7.9	-49.6	58.6	-2.32	0.0315
Smiley								
	DorsoLateral Prefrontal Cortex	BA 9	L	-15.5	45.7	38.6	6.10	0.0000
	Somatosensory and Motor Cortex	BA 3	R	33.6	-26.2	45.9	4.86	0.0001
	Temporo-Parieto-Occipital Junction	BA 39	L	-41.7	-69.5	11.8	4.11	0.0006
	DorsoLateral Prefrontal Cortex	BA 8	R	18.2	36.8	43.1	3.92	0.0009
	Early Visual Cortex	BA 19	L	-17.5	-73.4	-10.1	3.61	0.0018
	Somatosensory and Motor Cortex	BA 3	L	-38.6	-22.2	44.8	3.29	0.0038
	DorsoLateral Prefrontal Cortex	BA 10	R	27.8	46.1	14.7	2.42	0.0256
	Orbital and Polar Frontal Cortex	BA 10	L	-23.4	50.2	2.0	2.12	0.0471
Abstract								
	Early Visual Cortex	BA 19	L	-17.5	-73.4	-10.1	5.00	0.0001
	Somatosensory and Motor Cortex	BA 43	R	57.0	-8.6	18.6	5.02	0.0001
	Somatosensory and Motor Cortex	BA 3	R	33.6	-26.2	45.9	4.64	0.0002
	DorsoLateral Prefrontal Cortex	BA 9	L	-15.5	45.7	38.6	4.49	0.0003
	Superior Parietal Cortex	BA 2	L	-39.5	-34.8	35.4	3.81	0.0012
	DorsoLateral Prefrontal Cortex	BA 8	R	18.2	36.8	43.1	3.79	0.0012
	Anterior Cingulate and Medial Prefrontal Cortex	BA 24	L	-6.8	-6.4	40.2	3.56	0.0021
	Temporo-Parieto-Occipital Junction	BA 39	L	-41.7	-69.5	11.8	3.53	0.0022
	Paracentral Lobular and Mid Cingulate Cortex	BA 6	R	22.5	0.6	63.9	3.24	0.0044
	Inferior Frontal Cortex	BA 47	L	-48.8	30.1	-1.2	2.50	0.0216

Table 2: **P400 Maxima** The source maxima of at the P400 time window 300 ms - 600 ms after stimulus onset. HCP-MMP1: The Multi-Modal Parcellation atlas of the Human Connectome Project. BA: Brodmann Areas, X,Y,Z: MNI coordinates, T: T-value of the Wilcoxon test. p: corresponding uncorrected p-value of the Wilcoxon test.

located in ventral stream regions, particularly in the right IT; with no maxima located at the respective lateral temporal lobes (Fig. 5, Tab. 2). This activity may therefore be regarded as blur. We find further clusters of intersecting regions in the inferior frontal cortex (IFC), supplementary motor area (SMA), MCC and scattered bilateral regions in the parietal lobe. The LAURA inverse solutions reveal a similar pattern with intersecting activity in the IFC, bilateral IT/ Ventral stream regions, MC and inferior parietal lobe (IPL).

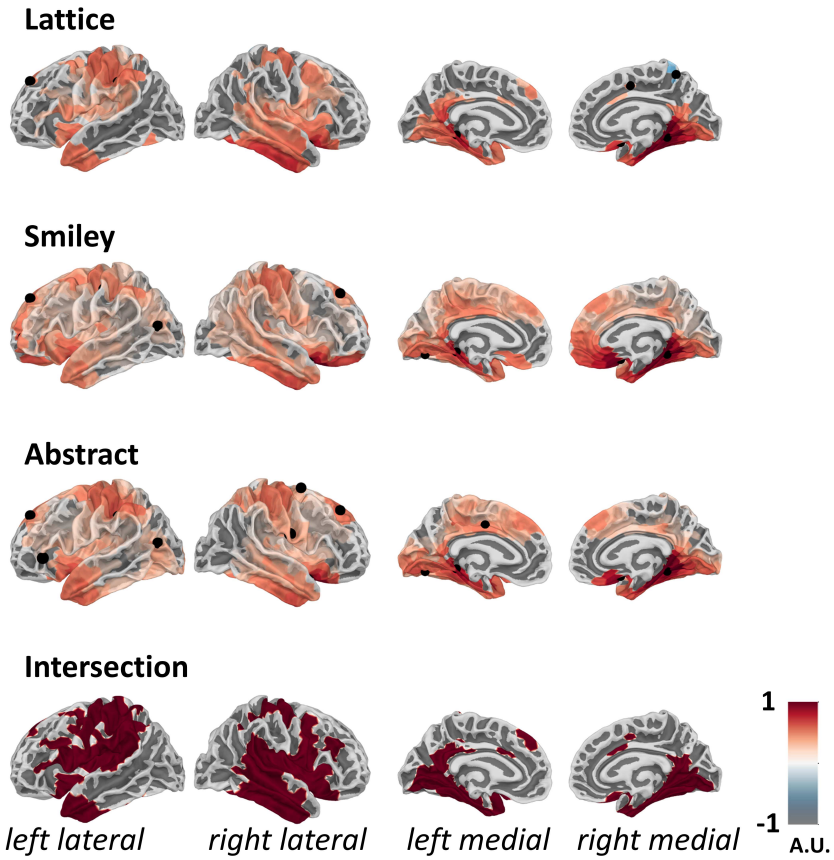


Figure 5: **FC-Sources of the P400** Grand average of the FC source differences unambiguous/high visibility - ambiguous/low visibility averaged across the P400 time range (300 - 600 ms post-stimulus) are shown. Units are arbitrary due to normalization of the individual differences. Only vertices that showed a significant difference at $\alpha < 0.05$ are shown. The intersection shows a binary mask of the vertices that were significantly active for all three stimuli. Sources in the left precentral gyrus, right medial temporal gyrus and right inferior temporal gyrus show larger activity for the unambiguous/ high-visibility stimuli throughout all categories.

5 Discussion

The human EEG shows a remarkable inter-individual variability that may be related to individual brain anatomies, conductivity and thickness of bones and meninges, etc. (e.g.,39). As one consequence, comparison of estimated brain sources based on EEG source localization suffers not only from the inverse problem but also from the inter-individual variability in brain anatomies. The latter problem can be reduced to some degree, if individualized brain models, based on anatomical MRI scans, are available and can be integrated into the source analysis algorithms. For the present study, no anatomical MRI scans were available. We thus decided to follow two parallel analysis traces:

(1) We focused on a within-participant statistics calculating a within-participant similarity analysis in order to estimate how similar the individually identified brain sources are across the three different categories of visual stimuli, given highly similar EEG topographies. This was based on correlations between the identified sources of the two conditions (ambiguous vs. disambiguated stimuli and low vs. high visibility stimuli).

(2) Additionally, we used the source space grand means across participants in order to estimate the shared brain sources of the ERP Uncertainty Effects across the stimulus categories.

For the within-participant analysis, we found highly significant correlations of estimated brain sources of the ERP Uncertainty Effects between smiley stimuli and abstract figures. We also found significant but less strong correlations between lattices and smiley on one hand and between lattices and abstract figures on the other hand. The grand mean analysis revealed a number of brain sources that were restricted to individual stimuli, but also brain sources that were strikingly similar across stimulus categories.

In the following we concentrate on those sources common to all stimuli: Using our ANN source localization approach, the P200 is mainly characterized by sources in the bilateral LOC, inferior parietal cortex and at the anterior cingulate cortex. The P400 is characterized

by bilateral superior parietal activations and deeper sources in the IT cortex, the latter with some lateralization to the right. These regions were also found with the LAURA inverse solutions, which serves as a methodological validation of the ANN.

Only few studies have contrasted neural activity underlying periods of stable percepts and compared these between ambiguous stimuli and disambiguated variants thereof. [16] compared BOLD responses to the Necker cube with BOLD responses from a disambiguated cube version. In this contrast they found a significant bilateral BOLD effect in the superior and inferior parietal lobe (SPL/IPL), which is also indicated in our data. Further they found activity in the premotor area and the medial frontal gyrus (MFG), the latter is a close neighbor of the anterior cingulate cortex we found. Although this was not explicitly stated in their manuscript, they also show a significant difference within the LOC. Thusly, there is significant overlap between the findings of [16] and our findings in parietal and occipital cortices.

The large majority of related fMRI and EEG source imaging studies focused on the event of a perceptual reversal (e.g., 40, 3). Reversal-related signatures reflect at least partly unstable brain states. The present analysis, in contrast, focuses explicitly on brain states during (temporally) stable percepts, thus stable brain states. One may postulate that perception of transiently stable ambiguous figures may still involve reversal-related brain networks. However, it is well possible that specific reversal-related activity may take place in brain areas that are not necessarily active during periods of perceptual stability.

This is highlighted by [40], who found significantly increased BOLD responses in the bilateral ventral occipital cortex (close to IT) and the bilateral IPL during perceptual reversals of Boring's ambiguous old/young woman [2] and Rubin's ambiguous face/ vase [41], compared to stable percepts. These findings may correspond to our findings in IT during the P400 window and the IPL activity in the P200 window. However, the conclusions need be drawn with caution due to the differences in stimuli.

Only a few studies on ambiguous stimuli estimated neural generators based on EEG data. [35] calculated the sources underlying the EEG during perceptual reversals of the Necker cube using LAURA, a linear solution to the EEG inverse problem. They focused on the sources underlying two reversal-related ERP components: The Reversal Negativity (RN, 230-280 ms after reversal onset) and the late positive component (LPC, 400-470 ms after reversal onset; "Parietal Positivity" in [42]). Both ERP components have been described earlier [43, 44].

[35] found that the RN component is mostly explained by sources in the bilateral IT, with a slight lateralization towards the right hemisphere. The LPC was localized in bilateral anterior IT and middle temporal gyrus as well as the bilateral superior parietal lobe. The sources found in IT correspond well to our own findings, both in location and lateralization towards the right hemisphere.

In summary, we found highly coherent neural sources of the ERP Uncertainty Effects across three different stimulus categories, indicating comparable neural processes. This was shown on the level of individual participants and at the group level using two vastly different approaches to solve the inverse problem. We found some remarkable overlap between the reported common sources and findings from the literature.

As stated in the introduction, the information available to our senses is *a priori* incomplete, noisy and to varying degrees ambiguous. Our perceptual system has to find the most probable and reliable perceptual interpretation in order to interact successfully with the environment. Of course, the quality of the sensory evidence varies considerably. Ambiguous stimuli, like the Necker cube or low-visibility stimuli, like the smileys and abstract figures are extremely challenging cases for the perceptual system. However, visual phenomena, like pareidolia (e.g., 45) indicate that our perceptual systems are trained to automatically produce results, independent of the quality of the sensory input. It is reasonable to postulate the existence of a meta-perceptual instance that evaluates perceptual outcomes at a post

perceptual processing step and indicates the reliability of perceptual interpretations (e.g., 7, 46). The ERP Uncertainty Effects may reflect outcomes of such evaluation processes, generalized across stimulus categories, with high amplitudes in the case of high reliability and vice versa. The common neural sources identified in the present study may reflect the underlying network. If this speculative interpretation is correct, it may allow interesting predictions: In the case of deficiency, such an evaluation system may indicate high reliability for erroneous perceptual interpretations, as for example described in certain cases of schizophrenia [47, 48]. If this were the case, we should assume an untypical pattern of the ERP Uncertainty Effects and underlying network activity. In the next step, we plan to test this hypothesis.

6 Data Availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

[1] L. A. Necker. Observations on some remarkable phenomenon which occurs in viewing a figure of a crystal or geometrical solid. *London and Edinburgh Philosophical Magazine and Journal of Science*, 3:329–337, 1832.

[2] Edwin G. Boring. A new ambiguous figure. *The American Journal of Psychology*, 1930.

[3] Jan Brascamp, Philipp Sterzer, Randolph Blake, and Tomas Knapen. Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annual review of psychology*, 69:77–103, 2018.

[4] Jürgen Kornmeier and Michael Bach. Ambiguous Figures – What Happens in the Brain When Perception Changes But Not the Stimulus. *Frontiers in Human Neuroscience*, 6, 2012. ISSN 1662-5161. doi: 10.3389/fnhum.2012.00051.

[5] Gerald M. Long and Thomas C. Toppino. Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological bulletin*, 130(5):748, 2004.

[6] Randolph Blake and Nikos K. Logothetis. Visual competition. *Nature Reviews Neuroscience*, 3(1):13–21, 2002.

[7] Jürgen Kornmeier and Michael Bach. Object perception: When our brain is impressed but we do not notice it. *Journal of Vision*, 9(1):7–7, 2009.

[8] Ellen Joos, Anne Giersch, Kriti Bhatia, Sven P. Heinrich, Ludger Tebartz van Elst, and Jürgen Kornmeier. Using the perceptual past to predict the perceptual future influences the perceived present–A novel ERP paradigm. *Plos one*, 15(9):e0237663, 2020.

[9] von Schiller, P. Stroboskopische Alternativbewegungen. *Psychologische Forschung*, 17: 179–214, 1993.

[10] Kriti Bhatia, Ellen Joos, and Jürgen Kornmeier. Ambiguity vs. Visibility: How the Perceptual System Responds to Uncertainty. *Master Thesis*, 2020.

[11] Ellen Joos, Anne Giersch, Lukas Hecker, Julia Schipp, Sven P. Heinrich, Ludger Tebartz van Elst, and Jürgen Kornmeier. Large EEG amplitude effects are highly similar across Necker cube, smiley, and abstract stimuli. *PloS one*, 15(5):e0232928, 2020.

[12] Lukas Hecker, Rebekka Rupprecht, Ludger Tebartz van Elst, and Juergen Kornmeier. ConvDip: A convolutional neural network for better M/EEG Source Imaging. *bioRxiv*, 2020.

[13] Lukas Hecker, Rebekka Rupprecht, Ludger Tebartz van Elst, and Jürgen Kornmeier. Long-Short Term Memory Networks for Electric Source Imaging with Distributed Dipole Models. Preprint, Neuroscience, April 2022.

[14] Roberto D. Pascual-Marqui. Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: Exact, zero error localization. *arXiv preprint arXiv:0710.3341*, 2007.

[15] Rolando Grave de Peralta Menendez, Micah M. Murray, Christoph M. Michel, Roberto Martuzzi, and Sara L. Gonzalez Andino. Electrical neuroimaging based on biophysical constraints. *NeuroImage*, 21(2):527–539, February 2004. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2003.09.051.

[16] Toshio Inui, Shigeki Tanaka, Tomohisa Okada, Sadahiko Nishizawa, Masahiro Katayama, and Junji Konishi. Neural substrates for depth perception of the Necker cube; a functional magnetic resonance imaging study in human subjects. *Neuroscience Letters*, 282(3):145–148, March 2000. ISSN 03043940. doi: 10.1016/S0304-3940(00)00899-5.

[17] Jürgen Kornmeier, Rike Wörner, and Michael Bach. Can i trust in what i see? EEG evidence for a cognitive evaluation of perceptual constructs. *Psychophysiology*, 53(10):1507–1523, 2016. ISSN 1469-8986. doi: 10.1111/psyp.12702.

[18] Jürgen Kornmeier and Michael Bach. Object perception: When our brain is impressed but we do not notice it. *Journal of Vision*, 9(1):7–7, January 2009. ISSN 1534-7362. doi: 10.1167/9.1.7.

[19] Michael Bach. The Freiburg Visual Acuity Test-automatic measurement of visual acuity. *Optometry and vision science*, 73(1):49–53, 1996.

[20] World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20):2191–2194, November 2013. ISSN 1538-3598. doi: 10.1001/jama.2013.281053.

[21] J. Kornmeier, S. P. Heinrich, H. Atmanspacher, and M. Bach. The reversing" Necker Wall"-A new paradigm with reversal entrainment reveals an early EEG correlate. In *Investigative Ophthalmology & Visual Science*, volume 42, pages S409–S409. ASSOC RESEARCH VISION OPHTHALMOLOGY INC 9650 ROCKVILLE PIKE, BETHESDA, MD . . . , 2001.

[22] M Woo, J Neider, and T Davis. OpenGL Programming Guide. The Official Guide to learning OpenGL. *Addison-Wesley*, Version 1.1(2), 1998.

[23] Jürgen Kornmeier, Werner Ehm, Heiko Bigalke, and Michael Bach. Discontinuous presentation of ambiguous figures: How interstimulus-interval durations affect reversal dynamics and ERPs. *Psychophysiology*, 44(4):552–560, 2007.

[24] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, and Lauri Parkkonen. MEG and EEG data analysis with MNE-Python. *Frontiers in neuroscience*, 7:267, 2013.

[25] Pierre Ablin, Jean-François Cardoso, and Alexandre Gramfort. Faster independent component analysis by preconditioning with Hessian approximations. *IEEE Transactions on Signal Processing*, 66(15):4040–4049, 2018.

[26] Pierre Ablin, Jean-François Cardoso, and Alexandre Gramfort. Faster ICA under orthogonal constraint. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4464–4468. IEEE, 2018.

[27] Mainak Jas, Denis Engemann, Federico Raimondo, Yousra Bekhti, and Alexandre Gramfort. Automated rejection and repair of bad trials in MEG/EEG. In *2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, pages 1–4. IEEE, 2016.

[28] Mainak Jas, Denis A. Engemann, Yousra Bekhti, Federico Raimondo, and Alexandre Gramfort. Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage*, 159:417–429, 2017.

[29] B. Fischl, M. I. Sereno, R. B. Tootell, and A. M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284, 1999. ISSN 1065-9471. doi: 10.1002/(sici)1097-0193(1999)8:4⟨272::aid-hbm10⟩3.0.co;2-4.

[30] Manfred Fuchs, Jörn Kastner, Michael Wagner, Susan Hawes, and John S. Ebersole. A standardized boundary element method volume conductor model. *Clinical Neurophysiology*, 113(5):702–712, 2002.

[31] Lukas Hecker, Rebekka Rupprecht, Ludger Tebartz Van Elst, and Jürgen Kornmeier. ConvDip: A Convolutional Neural Network for Better EEG Source Imaging. *Frontiers in Neuroscience*, 15, 2021. ISSN 1662-453X.

[32] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, and Matthieu Devin. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.

[33] François Chollet et al. Keras. 2015.

[34] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[35] Michael A. Pitts, Antígona Martínez, Clea Stalmaster, Janice L Nerger, and Steven A Hillyard. Neural generators of ERPs linked with Necker cube reversals. *Psychophysiology*, 46(4):694–702, July 2009. ISSN 0048-5772. doi: 10.1111/j.1469-8986.2009.00822.x.

[36] Roberta Grech, Tracey Cassar, Joseph Muscat, Kenneth P. Camilleri, Simon G. Fabri, Michalis Zervakis, Petros Xanthopoulos, Vangelis Sakkalis, and Bart Vanrumste. Review on solving the inverse problem in EEG source analysis. *Journal of neuroengineering and rehabilitation*, 5(1):25, 2008.

[37] Matthew F. Glasser, Timothy S. Coalson, Emma C. Robinson, Carl D. Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F. Beckmann, and Mark Jenkinson. A multi-modal parcellation of human cerebral cortex. *Nature*, 536 (7615):171–178, 2016.

[38] Korbinian Brodmann. *Vergleichende Lokalisationslehre Der Grosshirnrinde in Ihren Prinzipien Dargestellt Auf Grund Des Zellenbaues*. Barth, 1909.

[39] Paul L. Nunez and Ramesh Srinivasan. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, USA, 2006.

[40] A Kleinschmidt, C BÏchel, S Zeki, and R S J Frackowiak. Human brain activity during spontaneously reversing perception of ambiguous ®gures. page 8, 1998.

[41] Edgar Rubin. Synsoplevede figurer: Studier i psykologisk analyse. *Gyldendal, Nordisk forlag*, 1915.

[42] Juergen Kornmeier and Michael Bach. Evidence for early visual processing in perceptual disambiguation of ambiguous figures. *Journal of Vision*, 4(8):249–249, 2004.

[43] Jürgen Kornmeier, Michael Bach, and Harald Atmanspacher. CORRELATES OF PERCEPTIVE INSTABILITIES IN EVENT-RELATED POTENTIALS. *Interna-*

tional Journal of Bifurcation and Chaos, 14(02):727–736, February 2004. ISSN 0218-1274, 1793-6551. doi: 10.1142/S0218127404009430.

[44] Jürgen Kornmeier and Michael Bach. Bistable perception—along the processing chain from ambiguous visual input to a stable percept. *International Journal of Psychophysiology*, 62(2):345–349, 2006.

[45] Jürgen Kornmeier and Gerhard Mayer. The alien in the forest OR when temporal context dominates perception. *Perception*, 43(11):1270–1274, 2014.

[46] Pascal Mamassian. Visual confidence. *Annual Review of Vision Science*, 2(1):459–481, 2016.

[47] Pamela D. Butler, Steven M. Silverstein, and Steven C. Dakin. Visual perception and its impairment in schizophrenia. *Biological psychiatry*, 64(1):40–47, 2008.

[48] Veith Weilnhammer, Lukas Röd, Anna-Lena Eckert, Heiner Stuke, Andreas Heinz, and Philipp Sterzer. Psychotic Experiences in Schizophrenia and Sensitivity to Sensory Evidence. *Schizophrenia Bulletin*, 46(4):927–936, July 2020. ISSN 0586-7614, 1745-1701. doi: 10.1093/schbul/sbaa003.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf