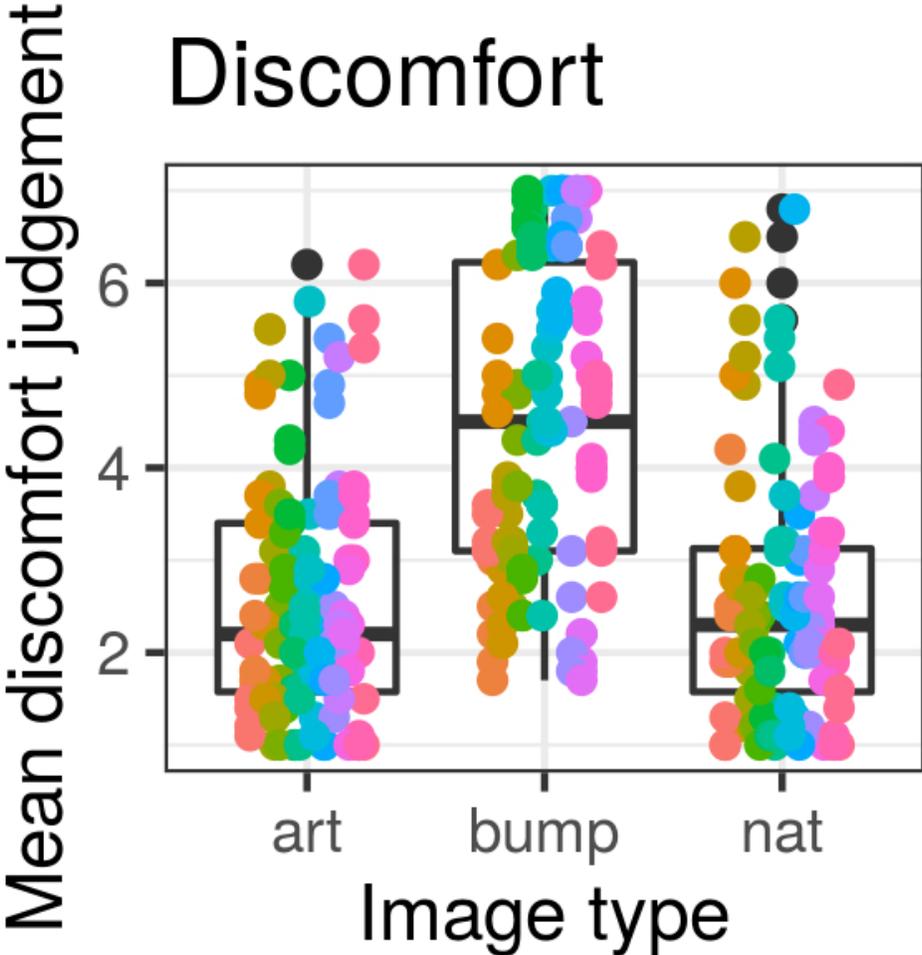


The visual system is thought to be optimised to process natural images efficiently. Abstract artworks may be aesthetically pleasing as they are more easily processed, conversely uncomfortable images may be less efficiently processed. When physical contrast is controlled for, fractal dimension and steady-state visual evoked potentials predict discomfort and aesthetic judgements of abstract artworks, natural and uncomfortable images, although the relationship is not as clear as efficient coding theories suggest.



Graphical Abstract: The visual system is thought to be optimised to process natural images efficiently. Abstract artworks may be aesthetically pleasing as they are more easily processed, conversely uncomfortable images may be less efficiently processed. When physical contrast is controlled for, fractal dimension and steady-state visual evoked potentials predict discomfort and aesthetic judgements of abstract artworks, natural and uncomfortable images, although the relationship is not as clear as efficient coding theories suggest.

50x50mm (300 x 300 DPI)

Title: SSVEP responses predict visual discomfort judgements

Running title: SSVEP responses predict visual discomfort judgements

Louise O'Hare^{1,2*}, Emily Hird¹, Mercedes Whybrow¹

¹School of Psychology, University of Lincoln, Brayford Pool, Lincoln, LN6 7TS

²Department of Psychology, Nottingham Trent University, 50 Shakespeare Street,
Nottingham, NG1 4QF

* corresponding author, LOHare@lincoln.ac.uk

Keywords: EEG, natural images, artworks, spectral slope, fractal dimension

Wordcount: 8477

Abstract

It has been suggested that aesthetically pleasing stimuli are processed efficiently by the visual system, whereas uncomfortable stimuli are processed inefficiently. This study consists of a series of three experiments investigating this idea using a range of images of abstract artworks, photographs of natural scenes, and computer-generated stimuli previously shown to be uncomfortable. Subjective judgements, and neural correlates were measured using EEG (steady-state visual evoked potentials, SSVEPs). In addition, global image statistics (contrast, Fourier amplitude spectral slope and fractal dimension) were taken into account. When effects of physical image contrast were controlled, fractal dimension predicted discomfort judgements, suggesting the SSVEP response is more likely to be influenced by distribution of

edges than the spectral slope. Importantly, when effects of physical contrast and fractal dimension were accounted for using linear mixed effects modelling, SSVEP responses predicted subjective judgements of images. Specifically, when stimuli were not matched for perceived contrast, there was a positive relationship between SSVEP responses and how pleasing a stimulus was judged to be, and conversely a negative relationship between discomfort and SSVEP response. This is significant as it shows that the neural responses in early visual areas contribute to the subjective (un)pleasantness of images, although the results of this study do not provide clear support for the theory of efficient coding as the cause of perceived pleasantness or discomfort of images, and so other explanations need to be considered.

Introduction

It has been suggested that the goal of the visual system is to process visual information efficiently (Barlow, 1961), especially given the high metabolic cost of the cortex (Lennie, 2003). In order to maximise efficiency, the human visual system is thought to be optimised for its environment (Field, 1987; Atick, 1992; Ruderman et al., 1998; Olshausen and Field, 2004). The natural visual environment is not random, but contains statistical regularities, such as a typical spatial frequency amplitude spectrum (Burton and Moorhead, 1987, Field, 1993; Tolhurst et al., 1992; Olmos and Kingdom, 2004). Cells in the early visual cortex have been suggested to respond in a way that is similar to Gabor filters (Daugman, 1980; Field, 1993), and computational studies have shown that Gabor filters would provide one possible method to transmit the information typically contained in natural images in a sparse and metabolically efficient manner (Field, 1987; Field, 1994; Field, 1999; Brady and Field, 2000).

Aesthetically pleasing images may further reduce processing redundancy (Ramachandran and Hirstein, 1999). Artworks may be aesthetically pleasing as they have spectral slope values even closer to the optimum for the visual system (Graham and Field 2007; Graham and Field 2008; Redies 2007; Mather, 2013; Koch et al., 2010). It has been suggested that artists deliberately adjust the spectral slope of the subject matter for artistic effect (Mather, 2014). Moreover, it has also been shown that preferences for faces are related to their Fourier spectral properties (Menzel et al., 2015).

Natural images and artworks have typical, and therefore predictable, statistical properties, and are generally considered to be aesthetically pleasing. By contrast, images with properties very different to those of natural images can be judged uncomfortable to look at (Fernandez and Wilkins, 2008; Juricevic et al., 2010; O'Hare and Hibbard, 2011). Computational models have demonstrated that discomfort from such stimuli could arise from inefficient, and therefore metabolically costly, visual processing (Pennachio and Wilkins, 2015; Hibbard and O'Hare, 2015). Evidence for this non-natural structure of uncomfortable images has also been shown in the temporal domain, in the form of flicker (e.g. Yoshimoto et al., 2017).

The amplitude spectrum is one predictable, statistical property of natural scenes. The alignment of local edges in an image (the phase spectra) of natural images also tend to have predictable properties that can be exploited for a sparse, efficient code (e.g. Field, 1993). The phase spectra of natural images tend to be statistically fractal patterns (Field, 1993), which are prevalent throughout natural scenes (Pentland, 1984). Although fractal patterns have amplitude spectra close to -1 (e.g. Field and Brady, 1997), not all images with $1/f$ amplitude spectra are fractal. A more direct measure is to estimate the fractal dimension by using methods such as “box-counting” (Li, 2009).

Fractal dimension influences judgements of aesthetics for both natural (Aks and Sprott, 1993; Spehar et al., 2003) and computer-generated images (Spehar et al., 2003). Works of abstract artwork, such as Pollock's drip paintings, have also been found to have a fractal structure (Taylor et al., 1999). In addition, Hagerhall et al., (2015) showed that fractal dimension influences the electrophysiological response of observers, measured using EEG.

Sarasso et al., (2020) showed using computer-generated images of varying slope exponent that the most aesthetically pleasing images were those with slopes of 1.8. The authors also showed a positive relationship between the C1 and N1 component amplitudes and aesthetic judgements. Additionally, fMRI evidence shows increasing activation in the occipital cortex in response to artworks that were judged to be more aesthetically pleasing compared to others (Vartian and Goel, 2004). This would be in conflict with the idea of efficient processing. It has been suggested that aesthetically pleasing stimuli might result in sensory enhancement in low-level visual areas could be due to the effects of attention (for a review, see Nadal, 2013). Stimuli with increased valence may recruit additional attentional resources and therefore result in an increased neural response. The idea of attentional enhancement is also supported with measures of task performance. Stimuli with spectral slope values of 1.8 were judged to be more aesthetically pleasing, and also showed reduced reaction times for a visual search task with these stimuli as a background (Sarasso et al., 2020). Additionally, for stimuli whose naturalness can be captured with low-level features, there seems to be a correspondence between faster reaction times and aesthetic judgements (Kardan et al., 2015).

There is a large body of theoretical and computational work, but there are fewer studies directly comparing neural responses to art, natural and uncomfortable images to assess the idea of efficient coding. One previous study using a limited sample of images of artworks, natural and uncomfortable images showed a tendency for images of artworks to have both lower discomfort judgements, and lower ERP responses compared to uncomfortable stimuli

and natural images (O'Hare and Goodwin, 2018). Steady-state visual evoked potentials (SSVEP) is an alternative method that involves a rapid presentation of stimuli ("flickering") for a period of time, for example 60 seconds, in order to evoke an oscillatory (steady-state) response, which has the benefit of an increased signal-to-noise ratio compared to other EEG methods (Vialette et al., 2010; Norcia et al., 2015). This means a wider selection of images can be included in the sample, increasing the robustness of the findings.

The visual system varies in contrast sensitivity (Campbell and Robson, 1968) and this can be seen in EEG responses (e.g. Plant, 1983, Regan, 1987). The stimuli previously used were matched for perceived contrast (O'Hare and Goodwin, 2018), therefore Experiment 1 will again use the same stimuli that have been matched. Experiment 2 will investigate stimuli matched for physical (root-mean-squared, RMS) contrast. Lastly, Experiment 3 includes stimuli chosen based on the spectral slope exponent.

Individual variation is a ubiquitous issue in work studying subjective judgements. Spehar et al., (2015) show that some elements of individual variation in preference could be related to the individual's sensitivity to that image attribute. Additionally, the relationship between fractal dimension and aesthetics judgements is not straightforward, and there is evidence of different subpopulations who show divergent patterns of responses – some preferring higher, others preferring lower fractal dimension values (Bies et al., 2016). As the relationship between subjective judgements and SSVEP responses is of interest, this relationship will be explored using linear mixed effects models. This allows for the random variation from individual differences to be included in the model, which is critical in this context, as well as including mediating effects such as low-level image statistics. In addition, as the interpretation of the specific word "discomfort" is unclear, observers will be asked to respond using a variety of slightly different phrasings, which will facilitate the generalisability of any consistent findings. If the efficient coding hypothesis is correct, the main hypotheses for the

current experiment are firstly, stimuli with statistical properties very different from those of natural images will be more uncomfortable than those with the statistical properties typical of natural images. Secondly, there will be a positive relationship between SSVEP response amplitude and the subjective discomfort of the images (conversely a negative relationship between how pleasing an image is and the amplitude of the SSVEP response).

Materials and Methods

Apparatus

Stimuli were presented using an MSI (MS-7788) computer with i7-3990CPU Intel processor, NVida GeForce GTX 650 graphics card. The display used was a 22" Illyama Vision Master Pro 514 monitor set to a resolution of 1024 x 768 with a refresh rate of 85Hz. The display was calibrated using a Minolta LS100 photometer, mean luminance was 44.24cd/m². A Bits # stimulus processor (Cambridge Research Systems) was used to convert the RGB signal to greyscale. Stimuli were presented using MATLAB (2015a) and the Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). EEG recordings were made using a 64-channel Active-Two Biosemi system and a 10-20 system electrode cap for the placement of the electrodes. There were eight facial electrodes, two on the mastoids, two on the outer canthi, two supraorbital and two infraorbital, in order to record eye movements.

Stimuli

Stimuli were artworks taken from the Guggenheim open collection (<https://www.guggenheim.org/collection-online>). These were one from each category of Bauhaus (Klee, *In der Strömung Sechs Schwellen*), Supremism (Malevitch, *Untitled*), Abstract Expressionism (Pollock, *Untitled (Green Silver)*), de Stijl (Mondrian, *Composition*

no VII) and Orphism (Kupka, Study for Amorpha, Warm Chromatic and for Fugue in Two Colours Study for the Fugue). Images were converted to greyscale using the Matlab function “`rgb2gray`”.

Bump images were the same as those used originally by (Fernandez and Wilkins, 2008), filtered noise patterns created using a raised radial cosine function (see equation 1).

$$H(f) = T \begin{cases} \frac{T}{2} [1 + \cos(\frac{\pi T}{\beta} (|\log(f) - \log(f_0)| - \frac{1-\beta}{2T}))] & 0 \leq |\log(f) - \log(f_0)| \leq \frac{1-\beta}{2T} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{for } \begin{cases} (0 \leq |\log(f) - \log(f_0)| \leq \frac{1-\beta}{2T}) \\ (\frac{1-\beta}{2T} \leq |\log(f) - \log(f_0)| \leq \frac{1+\beta}{2T}) \\ (|\log(f) - \log(f_0)| > \frac{1+\beta}{2T}) \end{cases}$$

Equation 1

Where T is 0.9, β is 0.5, f is the spatial frequency, and f_0 is the centre frequency of the function, defining the peak of the ‘bump’. The centre frequency was set to 0.75, 1.5, 3, 6, 9 cycles per degree. Natural images were five images chosen from the van Hateren image database (van Hateren and van der Schaaf, 1998). Fourier analysis was used to determine the spectral slope of the first 200 images of the van Hateren image database. The five images

with the spectral slope closest to 1 were chosen (image numbers 6, 65, 76, 84, 91). Stimuli were presented in a Gaussian-edged window. The radius of the window was 3.23° , and the Gaussian edge had a σ value of 0.2153° .

For the first experiment, a pilot experiment was conducted to match stimuli for perceived contrast prior to the main EEG recording experiment. Briefly, this experiment involved five participants in a two-alternative forced choice (2AFC) task using the method of constant stimuli. Psychometric functions were fitted to obtain a Point-of-Subjective-Equality (PSE), the 50% point, for the contrast of each stimulus. These settings were then used to match the stimuli for perceived contrast. Further details of the contrast matching task, including the data, can be seen in the supplementary information of O'Hare and Goodwin, (2018). There were five from each stimulus category which included: five images of artworks, five natural images (with slopes close to 1), and five bump stimuli of varying peak spatial frequency. Figure 1 shows examples of the stimuli used in the study. These 15 stimuli used in Experiment 1 had all been previously matched for perceived contrast in the pilot experiment.

*****figure 1 here *****

Observers and Ethical Statement

Twenty-six young observers with normal or corrected-to-normal vision took part in Experiment 1. EEG data from two observers were lost due to technical issues. All participants in all three studies gave their informed consent in writing at the time of participation. All experiments were conducted in accordance with the guidelines of the Declaration of Helsinki (2013) and were scrutinised by the University of Lincoln School of Psychology ethics committee, approval number: PSY1617334.

Procedure

Participants were seated in an electrically insulated, sound-attenuated room. The viewing distance was 1 meter and head movements were restricted using a chinrest. Participants were asked to remain as still as possible during the recording. A mid-grey screen was presented, with a black fixation cross in the centre 0.8613° in diameter. Stimuli were “flickered” by increasing and decreasing in contrast with a frequency of 5Hz. Each flickering stimulus was displayed for 5 seconds (Experiment 1) 20 seconds (Experiment 2), or 10 seconds (Experiment 3). After this a mid-grey screen appeared again, asking the observer to rate the stimuli from 1-7 for discomfort, and secondly for how aesthetically pleasing the stimulus appeared to be (Experiment 1). Stimuli were presented in random order, and each stimulus appeared 10 times (Experiment 1), 3 times (Experiment 2) and 6 times (Experiment 3). Therefore, the total presentation time for each individual stimulus in Experiment 1 was 50 seconds in total, 10 repetitions of a 5 second duration presentation time. For Experiments 2 and 3, individual stimuli were presented for a total of 60 seconds each, for Experiment 2 this resulted from 3 repetitions of a 20 second presentation time, and for Experiment 3 there were 6 repetitions of a 10 second presentation time.

For Experiments 1 there were three stimulus categories, artworks, bump stimuli and natural images, and each category consisted of 5 stimuli each (total 15 stimuli). For Experiment 2 there were four stimulus categories, artworks, bump stimuli, striped patterns and natural images. There were 10 artworks, 5 bump stimuli, 5 striped patterns, and 10 natural images (total 30 stimuli). For Experiment 3 there were 9 artworks, 5 bump stimuli, and 10 natural images (total 24 images).

Analysis

EEG data was analysed using EEGLAB (Delorme and Makeig, 2004). Data were re-referenced to the linked mastoids and filtered using a 0.1 to 40Hz band-pass FIR filter, to

remove drift and line noise. Bad channels were removed using the EEGLAB built-in procedure using a standard threshold and removing those outside of this value. Data were divided into 2.5 second (Experiment 1) or two-second (Experiments 2 and 3) epochs. Blinks were removed using a thresholding procedure, removing any epoch $\pm 150\mu\text{V}$. Eye movements were corrected using the Gratton-Coles (1983) correction procedure. This involved selecting the ocular channels for measuring eye movements, and calculating a propagation factor associated with eye movements, based on the following criteria – spikes in voltage in the ocular channels of 20mV for 200ms, and then the remaining channels are corrected for these values. This was performed using the EXG4 and EXG6 electrodes, which are adjacent to and below the right eye, respectively. The Gratton-Coles technique has the advantage of removing artefacts associated with eye movements whilst retaining much of the data. Figure 2 shows the scalp topography of the response for Experiment 1.

*****figure 2 here*****

Epochs were averaged over image. Spectral analysis was performed using FFT to estimate the peak amplitude of the SSVEP. The frequency spectrum for Experiment 1 can be seen in Figure 3, showing the peak response at the fundamental frequency, and the harmonics.

*****figure 3 here*****

SSVEP responses to the stimuli were pooled across electrodes recording over the early visual areas (O1, Oz, O2). The SSVEP response was defined as the peak response at the 5Hz fundamental frequency to the stimuli averaged across these electrodes. Epochs were grouped into three image categories and averaged: those containing all artworks, all bump stimuli, and all natural images for the main analysis. For the correlations with the subjective judgements,

each image was considered separately. The programme "R" (R Core Team, 2013) was used to analyse the data, using the package "afex" (Singmann et al., 2016), and quoting the generalised eta squared ($G\eta^2$) as an estimate of effect size. Greenhouse-Geisser correction was applied to adjust the degrees of freedom where assumptions of sphericity were not met. Least-squares means were reported as estimated group values, as well as model estimates for the contrasts between conditions with an indication of statistical significance at the level of $p < 0.05$. Linear mixed models were conducted using the package "lme4" (Bates et al., 2012). Bayes information criterion (BIC) was used to compare the strength of the evidence in favour of the linear mixed model in comparison to the null model, using the guidelines of Kass et al., (1995) to estimate the strength of the evidence. Generalisability of the model was estimated using cross-validation which is performed by splitting the data into a training set and test set, fitting the model on the training subset, and using this to predict the test (holdout) set (Song et al., 2021). This process is repeated to obtain an estimate of the generalisability of the model from the full data set (de Rooij and Weeda, 2020; Song et al., 2021). As the design had non-independent structure due to the repeated measures, cross-validation of the model was performed by removing data from one individual at a time to create the training and validation sets, (Colby and Blair, 2013).

Results

Subjective judgments

Figure 4 shows the effect of image type on discomfort judgments ($F(1.38,31.84) = 23.76$, $p = 4.86 \times 10^{-6}$, $G\eta^2 = 0.35$). Bump stimuli were rated as more uncomfortable compared to the other image categories, with a least-squares mean of 4.53 (SE = 0.26) compared to 2.57 (SE = 0.22) for artworks (estimate 1.95, SE = 0.34, $p = 4.0 \times 10^{-6}$) and 2.58 (SE = 0.22) for

natural images (estimate 1.94, SE = 0.40, $p = 0.0002$). There was no statistically significant difference between discomfort judgements for artworks and for natural images, with least-squares means of 2.57 and 2.58 respectively (estimate -0.01, SE = 0.20, $p = 0.99$).

*****figure 4 here*****

Figure 5 shows the effect of image type on judgements of how aesthetically pleasing the images were ($F(2.00, 46.00) = 64.47$, $p = 4.54 \times 10^{-14}$, $G\eta^2 = 0.61$). Artworks were judged to be more aesthetically pleasing than the other image types, with a least-squares mean of 4.65 (SE = 0.19) compared to 1.96 (SE = 0.16) for bump stimuli (estimate 2.69, SE = 0.24, $p = 3.25 \times 10^{-11}$) and 3.35 (SE = 0.21) for natural images (estimate 1.30, SE = 0.24, $p = 6.72 \times 10^{-6}$). Natural images were also considered an estimated 1.39 (SE = 0.24) times more aesthetically pleasing compared to bump stimuli, with least-squares means of 3.35 (SE = 0.21) and 1.96 (SE = 0.16) respectively ($p = 2.97 \times 10^{-6}$).

*****figure 5 here*****

SSVEP responses

Figure 6 shows the effect of image type on SSVEP responses ($F(1.55, 35.76) = 19.67$, $p = 8.2 \times 10^{-6}$, $G\eta^2 = 0.06$). The SSVEP response for natural images (least-squares mean of 55.7, SE = 0.52) was lower than for images of artworks (least-squares mean of 56.8, SE = 0.57) (estimate -1.04, SE = 0.27, $p = 0.002$) and also bump stimuli (least-squares mean of 57.3, SE = 0.51) (estimate = -1.60, SE = 0.31, $p = 0.0001$). The response to bump stimuli (least-squares mean of 57.3, SE = 0.53) was greater compared to images of artworks (least-squares mean 56.8, SE = 0.57) (estimate = 0.56, SE = 0.19, $p = 0.012$).

*****figure 6 here *****

Image statistics

Figure 7 shows the fractal dimension, spectral slope and RMS contrast of the three different image types. In this set of images, there was no statistically significant effect of image type on spectral slope ($F(2,12) = 0.002$, $p = 0.99$), on fractal dimension ($F(2,12) = 0.12$, $p = 0.89$), but there was a statistically significant effect of image type on RMS contrast ($F(2,12) = 3.95$, $p = 0.048$) in the current set of images. Least-squares means were lower for bump stimuli ($1.22 \pm 0.02SE$) compared to natural images ($0.20, \pm 0.02SE$, estimate = -0.08 , $SE = 0.03$, $p = 0.06$), and artworks ($0.19, \pm 0.02SE$, estimate = 0.07 , $SE = 0.03$, $p = 0.09$) although the difference between natural images ($0.20, \pm 0.02SE$) and artworks was much smaller (estimate = -0.01 , $SE = 0.03$, $p = 0.97$).

*****figure 7 here *****

Although image statistics are related to one another, there were only weak correlations between the image statistics of the stimuli used in Experiment 1: Spearman's rho (ρ) was used due to the non-normal distribution, and results showed only relatively weak correlations between spectral slope and fractal dimension ($\rho(13) = 0.31$, $p = 0.25$) and between spectral slope and RMS contrast ($\rho(13) = -0.23$, $p = 0.41$). There was no evidence of a relationship between fractal dimension and RMS contrast in the current set of images ($\rho(13) = -0.03$, $p = 0.91$). Correlations can be seen in figure 8.

*****figure 8 here*****

SSVEP response amplitude was predicted by spectral slope ($\chi^2(1) = 47.67$, $p = 5.04 \times 10^{-12}$), SSVEP response amplitude increased with spectral slope by $1.32 (\pm 0.18$ standard error). The

difference in BIC was 41.79, showing strong evidence in favour of the alternative model compared to the null model. This can be seen in Figure 9 (left). Cross validation showed the cross-validated $R^2 = 0.25$, suggesting 25% of the variance explained, and cross-validated MSE of 10.13 indicates that on average, model-fitted results will deviate from observed results by $\sqrt{10.13}$. SSVEP response amplitude was predicted by fractal dimension ($\chi^2(1) = 13.51$, $p = 0.0023$). SSVEP amplitude increased with fractal dimension by 46.92 (± 12.64 standard error), the difference in BIC was 7.62. This can be seen in Figure 9 (right). Cross-validation showed R^2 to be 0.18, suggesting 18% of the variance explained, and the cross-validated MSE was 10.49, and so on average, model-fitted scores will deviate from the observed results by $\sqrt{10.49}$. RMS contrast did not statistically significantly predict SSVEP response amplitude ($\chi^2(1) = 1.49$, $p = 0.22$).

*****figure 9 here*****

Relationship between subjective ratings and SSVEP responses

There were two linear mixed effects models conducted, one to predict discomfort judgements, and another to predict aesthetics judgements. Firstly, a linear mixed effect model was used to predict discomfort judgements. The model included SSVEP amplitude, spectral slope, fractal dimension and image type as fixed effects, and observer as a random intercept effect:

$$\text{discomfort} \sim \text{SSVEP} + \text{spectral slope} + \text{fractal dimension} + \text{RMS contrast} + \text{image type} + (1 | \text{observer}) + \text{error}$$

This full model was compared to a null model omitting SSVEP response amplitude. There was no statistically significant difference between the two models ($\chi^2(1) = 0.0005$, $p = 0.98$).

Similarly, there was also no statistically significant predictive effect of SSVEP response amplitude on judgements of how aesthetically pleasing the stimuli were ($\chi^2(1) = 0.0034$, $p = 0.95$).

As stimuli were matched for perceived contrast, but not physical (RMS) contrast, and contrast plays a role in discomfort judgements, the relationship between RMS contrast and discomfort was explored. RMS contrast predicted discomfort judgments ($\chi^2(1) = 45.80$, $p = 1.32 \times 10^{-11}$), contrast reducing discomfort judgements by $-9.59 (\pm 1.37 \text{ standard error})$. The difference in BIC was 39.91, suggesting strong support in favour of the alternative model. The cross-validated R^2 was 0.2, and so 20% of the variance was explained, and MSE was 2.73, suggesting that on average, model-predicted scores will deviate from the observed results by $\sqrt{2.73}$. Additionally, RMS contrast predicted aesthetics judgements ($\chi^2(1) = 51.99$, $p = 5.57 \times 10^{-13}$), contrast increasing aesthetics judgements by $10.00 (\pm 1.33 \text{ standard error})$ the difference in BIC being 46.11. Cross-validation showed R^2 to be 0.16, and cross-validated MSE to be 2.21. Figure 10 shows the bivariate correlations of the discomfort judgements and aesthetics judgements with SSVEP and each of the image statistics.

*****figure 10 here *****

In addition to the linear mixed effects models, the SSVEP response amplitude to the most and least uncomfortable images were compared using a t-test. The most and least uncomfortable images were determined by obtaining an average over all observers for each image. This was then split into two groups based on the scores higher or lower than the median value, respectively. Welch's two sample t-test was conducted due to unequal variances, which showed no statistically significant difference in the SSVEP response between the most and least uncomfortable images ($t(9.68) = 1.49$, $p = 0.17$). The same analysis was repeated for the aesthetics judgements, and again there was no statistically significant difference in the

SSVEP responses between the most and least aesthetically pleasing images ($t(11.44) = -0.42$, $p = 0.68$).

The SSVEP responses to the images with different fractal dimension values was also compared in the same manner. There was no statistically significant difference in SSVEP responses to images with higher or lower fractal dimension values ($t(11.92) = 2.10$, $p = 0.06$).

Interim Discussion

The results of Experiment 1 showed that, in a limited set of perceptually-contrast matched images, the low-level image properties of images all affected SSVEP responses measured over early visual areas. SSVEP responses are predicted by spectral slope and fractal dimension, but not by RMS contrast. SSVEP responses did not predict subjective judgements, whether discomfort or aesthetics. However, RMS contrast did predict discomfort and aesthetics judgements. In Experiment 2, SSVEP responses were measured again, with an expanded image set that were matched for physical (RMS) contrast.

Experiment 2

Experiment 2 was conducted to estimate the SSVEP when stimuli equalised for physical contrast, rather than perceptual contrast. Physical contrast was determined as the root-mean-squared (RMS) contrast of the stimuli. In addition, the study was to expand the range of artworks and natural images for comparison. This time the selection of natural images was chosen to avoid man-made structures, in order to have the stricter sense of "natural" images. In addition, vertical sine-wave stripe images were also included in the image set, as these have also previously been judged to be uncomfortable stimuli.

Method

General method was as Experiment 1. Data from eight observers were included in the analysis, as one observer could not be included due to technical issues. Stimuli consisted of artworks, bump stimuli, stripe stimuli and natural images. The artworks were nine images of artworks held at the Museum of Art Trento and Rovereto, which have been used in previous studies (Yanulevskaya et al., 2012). These were images “Studio per il ciclo Protesta” by Vedova; “Oggetto ottico dinamico” by Dadamaino; “Omaggio alla vecchia signora” by Scanavino; “Circolarità rotatoria” by Senesi; “Intermedio cromatico” by Senesi; “Kleine Welten VI” by Kandinskij; “Untitled” by Collins; “Erik Satie” by Veronesi; and “Multiplo” by Bonalumi. As before, images were converted to greyscale using the Matlab function "rgb2gray". The bump stimuli were as before, with peak spatial frequencies of 0.375, 0.75, 1.5, 3, 6cpd. Vertical sine-wave gratings were also included, with spatial frequencies of 0.375, 0.75, 1.5, 3, 6cpd. In addition, 10 natural images from the van Hateren image database (van Hateren and van der Schaaf, 1998) were included, this time a list of all those images without man-made features was compiled by visual inspection. From this list, 10 images were chosen at random, the final image numbers included were 16, 31, 37, 53, 152, 155, 156, 190, 192, 196. This time, stimuli were matched for RMS contrast. Stimuli were presented three times in total, for 20 seconds each, with a fixation cross presented in the centre of the screen. Observers were asked to rate how pleasing the images were.

Results

Data were analysed as in Experiment 1.

The effect of image type on how pleasing images were judged to be ($F(1.91,13.37) = 5.81$, $p = 0.002$, $G\eta^2 = 0.31$) can be seen in Figure 11. Natural images (least-squares mean = 4.46, SE = 0.22) were considered to be more pleasing compared to artworks (least-squares mean = 3.47, SE = 0.32, estimate = 0.99, SE = 0.29, $p = 0.04$), bump stimuli (least-squares mean = 2.68, SE = 0.55, estimate = 1.78, SE = 0.40, $p = 0.01$), and stripe images (least-squares mean = 2.74, SE = 0.43, estimate = 1.72, SE = 0.53, $p = 0.05$). Artworks (least-squares means = 3.47, SE = 3.2) were considered to be more pleasing compared to bump stimuli (least-squares means = 2.68, SE = 0.55, estimate = 0.79, SE = 0.45, $p = 0.37$) and stripes (least-squares means = 2.74, SE = 0.43, estimate = 0.73, SE = 0.48, $p = 0.47$), although these comparisons were not statistically significant. Stripe images (least-squares means = 2.74, SE = 0.44) were not statistically significantly more pleasing compared to bump stimuli (least-squares means = 2.68, SE = 0.55, estimate = 0.05, SE = 0.69, $p = 0.99$).

*****figure 11 here *****

There was a statistically significant effect of image type on SSVEP response ($F(1.49,10.28) = 5.13$, $p = 0.036$, $G\eta^2 = 0.10$). SSVEPs to artworks (least-squares mean = 58.1, SE = 0.52) were greater than to bump stimuli (least-squares mean = 56.7, SE = 0.42; estimate = 1.39, SE = 0.21, $p = 0.001$), and natural images (least-squares mean = 57.5, SE = 0.50; estimate = 0.66, SE = 0.20, $p = 0.04$), but not statistically significantly different compared to stripes (least-squares mean = 57.5, SE = 0.73; estimate = 0.68, SE = 0.48, $p = 0.53$). SSVEPs to natural images (least-squares mean = 57.5, SE = 0.50) were greater compared to bump stimuli (least-squares mean = 56.7, SE = 0.42; estimate = 0.73, SE = 0.18, $p = 0.02$), and stripes (least-squares mean = 57.5, SE = 0.73; estimate = 0.02, SE = 0.41, $p = 1.00$), although this was not statistically significant in the current sample. There was no statistically

significant difference between stripes (least-squares mean = 57.5, SE = 0.73) and bump stimuli (least-squares mean = 56.7, SE = 0.42; estimate = 0.71, SE = 0.48, $p = 0.50$) in the current set of images. This can be seen in Figure 12.

*****figure 12 here *****

Figure 13 shows the image statistics for each of the image categories in Experiment 2. There was a significant effect of image type on spectral slope ($F(2,22) = 11.44$, $p = 3.39 \times 10^{-4}$). Least-squares means showed natural images (-0.15 , ± 0.17 SE) had a shallower slope compared to artworks (-1.05 ± 0.17 SE, estimate = -0.90 , SE = 0.23, $p = 0.002$) and bump stimuli (-1.33 , ± 0.23 SE, estimate = -1.18 , SE = 0.29, $p = 0.001$). There was no statistically significant difference between artworks (-1.05 ± 0.17 SE) and bump stimuli (-1.33 , ± 0.23 SE, estimate = 0.28, SE = 0.29, $p = 0.61$). Spectral slope was not included for sine waves as this is not possible to calculate. There was also a significant effect for fractal dimension ($F(3,26) = 7.55$, $p = 8.58 \times 10^{-4}$). Least-squares means were lower for artworks (1.83 , ± 0.01 SE) compared to bump stimuli (1.85 ± 0.01 SE, estimate = -0.02 , SE = 0.01, $p = 0.04$), stripes (1.86 , ± 0.01 SE, estimate = -0.03 , SE = 0.01, $p = 0.02$), and natural images (1.86 , ± 0.01 SE, estimate = -0.03 , SE = 0.01, $p = 0.0008$). There was no statistically significant difference between bumps (1.85 ± 0.01 SE) and stripes (1.86 , ± 0.01 SE, estimate = -0.002 , SE = 0.01, $p = 0.99$) or natural images (1.86 , ± 0.01 SE, estimate = -0.006 , SE = 0.01, $p = 0.86$), or between stripes (1.86 , ± 0.01 SE) and natural images (1.86 , ± 0.01 SE, estimate = -0.004 , SE = 0.01, $p = 0.97$). Stimuli were matched for RMS contrast.

*****figure 13 here *****

There was a positive relationship between spectral slope and fractal dimension in the stimulus set used in Experiment 2 ($\rho(28) = 0.35$, $p = 0.08$), see Figure 14.

*****figure 14 here *****

It was not possible to obtain spectral slope values for the sine waves, as these consist of only one frequency, therefore a reduced dataset excluding sine waves was created to investigate the relationship between spectral slope and SSVEP response amplitude. Image spectral slope does not statistically significantly predict SSVEP response amplitude ($\chi^2(1) = 0.01$, $p = 0.90$). Fractal dimension predicts SSVEP response amplitude ($\chi^2(1) = 42.87$, $p = 5.83 \times 10^{-11}$), reducing SSVEP by $-81.92 (\pm 11.94 \text{ standard error})$. The difference in BIC was 37.40, suggesting strong evidence in favour of the alternative model. Cross-validation R^2 was 0.19, suggesting 19% of the variance was explained, and MSE was 8.89, showing that on average, the fitted values deviate from the observed values by $\sqrt{8.89}$. This can be seen in Figure 15. RMS contrast was matched in these stimuli.

*****figure 15 here *****

A linear mixed effect model was used to investigate the predictive effect of SSVEP response amplitude on how pleasing an image was thought to be. As fractal dimension was predictive of subjective judgements, this was also included in the model:

$$\text{pleasing} \sim \text{SSVEP} + \text{fractal dimension} + \text{image type} + (1|\text{observer}) + \text{error}$$

There was a statistically significant prediction of SSVEP on pleasing judgements ($\chi^2(1) = 15.27$, $p = 9.31 \times 10^{-5}$), increasing by $0.17 (\pm 0.046 \text{ standard error})$. Fractal dimension reduced pleasing judgements by $-1.41 (\pm 6.56 \text{ standard error})$. Cross-validation showed R^2 to be 0.33 and MSE to be 1.63. The difference in BIC was 9.79 suggesting strong evidence in favour of the alternative model (Kass et al., 1995). Figure 16 shows the bivariate correlations

for judgements of how pleasing a stimulus was perceived to be against SSVEP, and the image statistics.

*****figure 16 here*****

Again, SSVEP responses to most and least pleasing stimuli, determined by median split, were compared. Welch's two-sample t-test showed significantly higher SSVEP responses for images that were judged to be more pleasing (mean = 58.16, SD = 2.61) compared to those that were less pleasing (mean = 56.53, SD = 1.61), $t(23.33) = 2.63$, $p = 0.015$.

There was no statistically significant difference in the SSVEP response to images with higher or lower fractal dimension values ($t(25.39) = -1.22$, $p = 0.15$).

Interim Discussion

Images matched for physical (RMS) contrast showed only a relatively small effect of image type on SSVEP responses, the only difference being between artworks and bump images. Importantly, when stimuli were matched for RMS contrast, there was no effect of spectral slope on the SSVEP response, only fractal dimension. However, SSVEP response amplitude in this stimulus set was predictive of how pleasing they were, the greater the SSVEP amplitude, the more pleasing the image was judged to be. In addition, the higher the fractal dimension, the less pleasing the image was judged to be. The combined results of Experiment 1 and 2 show that when the effects of physical contrast are controlled for, subjective judgements can be predicted by SSVEP responses. Although spectral slope and fractal dimension are different, they are relatively closely related image statistics. To separate their contributions, the images for Experiment 3 were chosen to minimise the difference in spectral slope values between artworks and bump stimuli, whilst allowing fractal dimension to vary.

Experiment 3

The stimuli in Experiment 3 were chosen based on their spectral slope values, to reduce the differences in average spectral slope values between artworks and natural image categories.

Method

General method was as Experiment 1. 21 young observers took part in the study. Spectral slope exponent, fractal dimension and RMS contrast were estimated for all stimuli in the MART stimulus set. The artworks were chosen based on having spectral slopes closest to 1 out of the MART selection. This resulted in the following images being chosen: “Bandiera” by Bianco; “Concetto spaziale” by Fontana; “Teorema” by Belli; “Obliquo grigio” by Griffa; “Verticale” by Griffa; “Costruzione SN 6” by Veronesi; “Ciclo 61/62 n. 8” by Vedova; “Gammature di colore” by Marchegiani; “Achrome” by Manzoni; and “Il movimento delle cose” by Dadamaino. Stimuli from the natural images were the following images, 100, 111, 118, 124, 132, 155, 163, 175, 179, 198, chosen to match the artwork stimuli for spectral slope as closely as possible. Using an independent t-test, the difference between the spectral slope values for the images of artworks and natural images was marginally significant ($t(249.38) = -1.98, p = 0.049$). Bump stimuli were as before. Figure 17 shows the image statistics included in the final set. There was no statistically significant effect of image type on spectral slope ($F(2,22) = 0.026, p = 0.97$) or fractal dimension ($F(2,22) = 2.59, p = 0.10$), but there was a statistically significant effect of image type on RMS contrast ($F(2,22) = 4.45, p = 0.024$). Least-squares means showed there to be greater RMS contrast for natural images (0.18 ± 0.02) compared to artworks ($0.09 \pm 0.02SE$), but no statistically significant difference between the artworks ($0.09 \pm 0.02SE$) and bump stimuli ($0.12 \pm 0.03SE$), or bump stimuli ($0.12 \pm 0.03SE$) and natural images (0.18 ± 0.02). This time observers were asked to rate stimuli on a scale of 1-5 for how uncomfortable they found the stimuli.

*****figure 17 here *****

Results

Data were analysed as Experiment 1. Data from two observers were excluded from the analysis as there was insufficient data remaining after artefact rejection procedures.

Figure 18 shows a statistically significant effect of image type on discomfort judgements ($F(1.97,35.54) = 30.91, p = 1.85 \times 10^{-8}, G\eta^2 = 0.39$). Bump stimuli (least-squares mean = 3.11, SE = 0.21) were more uncomfortable than artworks (least-squares mean = 2.49, SE = 0.23; estimate = 0.63, SE = 0.21, $p = 0.002$), and natural images (least-squares mean = 1.56, SE = 0.12; estimate = 1.56, SE = 0.20, $p = 9.24 \times 10^{-8}$). Artworks (least-squares mean = 2.49, SE = 0.23) were considered more uncomfortable compared to natural images (least-squares mean = 1.56, SE = 0.12; estimate = 0.93, SE = 0.20, $p = 0.0005$).

*****figure 18 here *****

There was no statistically significant effect of image type on SSVEP response amplitude ($F(1.35,24.27) = 1.15, p = 0.31, G\eta^2 = 0.007$). This can be seen in Figure 19.

*****figure 19 here *****

Spectral slope was not completely matched, so a linear mixed effect model was used to estimate whether there was a predictive effect of spectral slope on SSVEP amplitude. No statistically significant effect was found ($\chi^2(1) = 0.03, p = 0.87$). There was no statistically significant effect of RMS contrast on SSVEP response ($\chi^2(1) = 3.39, p = 0.07$). However, once again, fractal dimension predicts SSVEP response amplitude ($\chi^2(1) = 7.47, p = 0.006$),

reducing SSVEP by $-14.30 (\pm 5.20 \text{ standard error})$. The difference in BIC was 1.35 in favour of the alternative model. Cross validation showed the R^2 to be 0.09 and the MSE to be 9.27, and so the model does not generalise. This can be seen in Figure 20.

*****figure 20 here *****

For the stimuli used in Experiment 3, there was only a weak relationship between spectral slope and fractal dimension ($\rho(23) = -0.18, p = 0.38$), or between spectral slope and RMS contrast ($\rho(23) = -0.11, p = 0.61$). However, there was a stronger positive relationship between fractal dimension and RMS contrast ($\rho(23) = 0.60, p = 0.0016$), see Figure 21.

*****figure 21 here *****

A linear mixed effect model was created to predict discomfort judgements, including SSVEP and image type, fractal dimension and RMS contrast as fixed effects, and observer as a random intercept effect:

$$\text{discomfort} \sim \text{SSVEP} + \text{fractal dimension} + \text{RMS} + \text{image type} + (1|\text{observer}) + \text{error}$$

This was compared to a null model without SSVEP. There was a statistically significant effect of SSVEP amplitude in predicting discomfort judgements ($\chi^2(1) = 6.04, p = 0.014$). Discomfort reduced with SSVEP amplitude by $-0.04 (\pm 0.012 \text{ standard error})$. The difference in BIC of -0.12 suggests there is no evidence in favour of the alternative model (Kass et al., 1995). Cross-validation showed there to be R^2 to be 0.56 and MSE to be 0.41, and so the model does generalise well. Figure 22 shows the bivariate correlations of the predictors included in the model with discomfort judgements.

*****figure 22 here *****

SSVEP responses to more and less uncomfortable images were determined by median split, and there was no statistically significant difference between these images according to Welch's two-sample t-test ($t(16.87) = -0.92, p = 0.37$). Similarly, there was no statistically significant difference in the SSVEP response to images of higher or lower fractal dimension ($t(19.18) = -1.88, p = 0.075$).

Interim Discussion

In images matched for spectral slope, the greater the SSVEP, the lower the discomfort judgements. This complements Experiment 2, where SSVEP increased judgements of how pleasing a stimulus was. RMS and perceived contrast were both free to vary in this experiment; natural images in this set had increased fractal dimension, but lower RMS contrast, and artworks had a lower fractal dimension, but increased RMS contrast. As in Experiment 1, there was no predictive effect of RMS contrast on SSVEPs in this experiment, however, both fractal dimension and RMS contrast were included in the model to control for their effects in predicting discomfort judgements. In this set of images, artworks were considered more uncomfortable than natural images, whereas in Experiment 1, no difference was seen between artworks and natural images. In both Experiment 1 and Experiment 3, natural images had the highest physical contrast value, but in Experiment 3 this difference was much bigger, and so this could account for the different findings.

General discussion

Subjective judgements

This study directly compared artworks, uncomfortable images, and natural scenes, measuring both image statistics, subjective judgements (using a variety of phrasings, in order to

generalise findings) and occipital SSVEP responses, in three separate experiments with three different sets of observers. Unnatural images were consistently judged as the most uncomfortable/least pleasing across all three experiments, supporting previous research (e.g. Fernandez and Wilkins, 2008; O'Hare and Hibbard, 2011). Discomfort judgements were influenced by physical RMS contrast in both Experiment 1 and Experiment 3. However, both SSVEP responses and fractal dimension can predict subjective judgements, once the influence of contrast has been reduced.

Spectral slope

Overall, spectral slope did not predict discomfort judgements. Other researchers have also found there to be only a weak relationship between spectral slope and discomfort judgements (Ogawa and Motoyoshi, 2020). Estimates of spectral slope should reflect some aspect of contrast gradients in the image as a whole (Field and Brady, 1997). However, it is important for the application of global image statistics, such as spectral slope, that the image statistics do not vary substantially on a local level (Field, 1999). Whilst this constraint is true for natural images (Olshausen and Field, 2000), it may not be the case for abstract artworks, which are less constrained by subject matter. As spectral slope only indirectly captures some aspects of local contrast gradients (edges) it may not reflect certain local image statistics important for discomfort. For example, blurred images have differences in spectral slope and are also considered uncomfortable (e.g. O'Hare and Hibbard, 2013). However, it must be noted that background blurring (for example depth-of-field blurring) is well-tolerated by observers (e.g. O'Hare et al., 2013) and can even improve judgements of aesthetic appeal in photographs (e.g. Datta et al., 2006). Therefore, this general image statistic may miss important information regarding discomfort.

Spectral slope has been shown to relate to EEG responses to computer-generated stimuli, specifically, greater ERP amplitudes for images with slope exponents around 1.8, which were also considered the most aesthetically pleasing (Sarasso et al., 2020). Spectral slope values were predictive of SSVEP responses when stimuli were matched for perceptual contrast (Experiment 1), but not when they were matched for RMS contrast (Experiment 2).

Therefore, it is possible that the relationship between SSVEP response amplitude and spectral slope depends on contrast.

Fractal dimension

In the current study, fractal dimension, but not spectral slope, was predictive of SSVEP amplitude in both Experiment 2 and Experiment 3. Additionally, fractal dimension was predictive of how pleasing an image was judged to be (Experiment 2). This indicates the importance of the distribution of edges in the scene. Fractal dimension reflects a different global image property that takes account some aspect of edge information, giving an estimate of "complexity" of the image. Complexity of representational artwork and natural scenes may well evoke feedback from higher visual areas, which has effects on neural responses from early visual areas (Murray et al., 2002; Bar et al., 2006). This could possibly relate to semantic content, but as fractal dimension is associated with art and aesthetics, even in abstract patterns (e.g. Taylor et al., 1999; Sephar et al., 2003), this would appear to be independent of image semantics. Increasing fractal dimension of abstract patterns also increases excitement judgements, and so the SSVEP response may be related to increased excitement, rather than discomfort (Abboushi et al., 2019).

There was also a relationship between SSVEP responses and fractal dimension for Experiment 1, this time a positive relationship, although weaker. It is unclear why this might be. Speculatively, the stimuli in Experiment 1 were matched for perceived contrast, and so

speculatively, this may have had an impact on the results, as fractal dimension is related to image contrast, specifically pixel variance (Shamsyeh Zahedi and Zeil, 2018). This can also be seen in the current study, when RMS contrast was controlled (Experiment 2) there was a positive relationship between spectral slope and fractal dimension – the higher the spatial frequency content, the more complex the image. When spectral slope was controlled (Experiment 3) there was a relationship between fractal dimension and RMS contrast – the more complex the image, the greater the physical contrast. SSVEP responses in Experiment 1 were also related to spectral slope, as well as fractal dimension, which could indicate a possible role of contrast, which is a common factor in these two metrics. From the results of this study, it is not possible to draw firm conclusions about the reasons for this difference in direction, it might be possible to investigate this more thoroughly in future research, by creating artificial stimuli that vary in fractal dimension in a more controlled fashion.

The importance of phase information

Phase information, and the exact location of edges, is important to visual discomfort. It has been suggested that text poses a challenging stimulus for the visual system due to its "striped" nature (Wilkins et al., 2007). In addition, op-art artworks, or op-art-based visual stimuli, that tend to consist of geometric patterns including stripes, have been shown to elicit illusory motion in observers (Wade, 1978; Troncoso et al., 2008; Zanker et al., 2003; Zanker, 2004; Zanker et al., 2010, Hermens et al., 2012), which is included in some definitions of visual discomfort (e.g. Wilkins et al., 1984). Aside from discomfort, Grebenkina et al., (2018) showed there to be effects on aesthetics judgments that could be predicted using edge-orientation entropy. Phase-scrambled artworks show reduced EEG responses compared to the originals (Menzel et al., 2018). Therefore, although the amplitude spectrum is important to discomfort when the phase spectrum is random (e.g. Fernandez and Wilkins, 2008; Juricevic et al., 2010; O'Hare and Hibbard, 2011), when considering cross-category comparisons, the

phase spectrum may be more important. The exact location and perceived sharpness of edges in an image will contribute to subjective judgements, which are not entirely accounted for by RMS contrast or spectral slope but may be better captured by fractal dimension.

Ecologically valid stimuli

The choice of using images of real artworks made the stimulus set relatively uncontrolled and less systematic than would have been possible by investigating purely computer-generated images. However, there is a need to use ecologically valid stimuli in research (Olshausen and Field, 2004), especially given the work by authors such as McManus et al., (1993), who showed that observers can reliably discriminate between real paintings by Mondrian and computer-generated ones in this style. This suggests that there are additional elements of image composition determining aesthetic judgements, which could not necessarily be recreated. Phase information relates to the semantic content of the image and is important in perception of visual stimuli (Oppenheim and Lim, 1981; Thomson 1999), and it is not clear which aspects of the phase spectrum is important. To investigate the phase spectra of images, it is possible to scramble the phase spectra of images (e.g. Dakin et al., 2002). However, it is much more complex to meaningfully and systematically vary this parameter. It would be possible to use wavelet filtering to systematically manipulate the amplitude spectrum of images, which has been used in the past (for a review, Simoncelli and Olshausen, 2001). However, there may be an unnatural quality imposed on images manipulated in this way, for example, images that have the spectral slope manipulated to reduce content at higher spatial frequencies can appear blurred, and this can itself affect discomfort judgements (O'Hare and Hibbard, 2013). Whitened images, those with an increase of higher spatial frequency content, can also look unnatural (Párraga et al., 2000). However, it is currently unclear how to systematically manipulate the fractal dimension of natural images at the present time. Using computer-generated images would allow these manipulations, however artificial images are

not the same as natural images, as outlined above. Therefore, the choice was made to use the less controlled, but more ecologically valid stimuli, relying instead on natural variation in phase spectra between the images. However, it would be beneficial in future research to investigate systematically varying the amplitude and phase spectra of natural images, using techniques such as wavelet filtering, or to create images of varying fractal dimension.

Relationship between SSVEP and subjective judgements

SSVEP responses were predictive of subjective judgements - there was a greater SSVEP response to pleasing stimuli (Experiment 2) and a reduced SSVEP response to uncomfortable stimuli (Experiment 3). At first glance, this appears to contradict the idea of efficient coding, as one would expect a reduced response to stimuli that were easier to process. An alternative explanation is that of attentional enhancement. Previous work has shown that aesthetically pleasing stimuli result in greater electrophysiological (e.g. Sarasso et al., 2020) and BOLD responses (e.g. Vartian and Goel, 2004) over the low-level visual areas, which is thought to be due to increased attentional processing of these more pleasing visual stimuli (Nadal, 2013). These studies have demonstrated this positive relationship between aesthetics and brain activity within stimulus categories, the results of the current study could be interpreted as supporting this idea *between* stimulus categories - it could be the case that in this study the most aesthetically pleasing stimulus categories were also those with greater attentional resources allocated. It might be possible to test the allocation of attentional resources to aesthetically pleasing stimuli using methods such as eye tracking (Nayak and Karmakar, 2019; Locher, 2020).

However, recent work comparing SSVEP responses at 6Hz showed reduced SSVEP amplitude to unpleasant stimuli compared to neutral stimuli (Bekhtereva and Müller, 2015; and correction to original paper, Bekhtereva and Müller 2017). However, the opposite result

(an increase in SSVEP response to unpleasant stimuli compared to neutral) was found for different SSVEP driving frequencies of 3Hz, 4Hz, and 8.57Hz (Bekhtereva et al., 2018). The authors provided modelling work to support their suggestion that this paradoxical effect is due to a linear superposition of the transient ERP components that cause destructive interference in this particular range (5-6Hz) of stimulation frequency, which does not occur at the other frequencies used (Bekhtereva et al., 2018). The idea that the SSVEP signal may be paradoxically reduced, rather than increased, also occurs when stimuli are presented to one or two eyes – Wu, (2016) showed that the SSVEP response to one eye was greater than to both eyes. It might be expected that the stimulation of both eyes would lead to a greater response, as more of the network is involved in processing. However, as part of the network is shared in the case of the two-eye stimulation, and there is a limit to the response rate (due to the “axon hillock”), then the overlapping regions respond alternately, and so the signal is reduced. Therefore the results of Wu, (2016) and Bekhtereva et al., (2018) suggest that the direction of the relationship may be more complicated than expected due to other factors such as destructive interference due to this particular choice of stimulation frequency. It would be necessary to conduct the experiment at different flicker frequencies in order to reconcile these findings with the previous ERP data on visual discomfort. Until this has been demonstrated empirically, alternative explanations such as attentional enhancement may provide a more parsimonious explanation.

Conclusion

As predicted by efficient coding, stimuli without the statistical properties of natural scenes were found to be more uncomfortable, less pleasing, and less aesthetically pleasing compared to natural images and images of artworks. Fractal dimension, but not spectral slope, predicted subjective judgements when physical contrast was controlled for, demonstrating the importance of the phase spectrum when comparing images across categories. When

accounting for low-level image properties, there was a relationship between subjective judgements and SSVEP responses, although this is not in the correct direction to support efficient coding, and so it is possible that this reflects other processes, such as attentional enhancement instead. This provides evidence that responses in early visual areas contribute to judgements of how pleasing/uncomfortable an image is judged to be.

Acknowledgements

The authors wish to thank Katie Ascott for assistance in data collection in Experiment 1.

There was no financial support for this research.

Competing Interests

None of the authors have any competing interests.

Author Contributions

LOH designed Experiment 1, contributed to data collection for Experiments 1, 2, and 3, analysed the data for all three experiments, and drafted the manuscript. MW contributed to the design and data collection for Experiment 2, and also in drafting the manuscript. EH contributed to the design and data collection in Experiment 3, and assisted in drafting the manuscript.

Data Accessibility

Data can be found at Figshare.com: 10.6084/m9.figshare.9285065

Abbreviations

2AFC - two-alternative forced-choice

BIC - Bayes information criterion

EEG - Electroencephalogram

ERP - Event-related potential

RMS - Root-mean-squared

SSVEP - Steady-state visual evoked potential

References

Abboushi, B., Elzeyadi, I., Taylor, R., Sereno, M., (2019). Fractals in architecture: The visual interest, preference, and mood response to projected fractal light patterns in interior spaces.

Journal of Environmental Psychology, **61**, 57-70. Doi:

<https://doi.org/10.1016/j.jenvp.2018.12.005>

Aks, D., and Sprott, J. C., (1996). Quantifying aesthetic preference for chaotic patterns.

Empirical Studies of the Arts, **14**, 1-16. Doi: 10.2190/6V31-7M9R-T9L5-CDG9

Atick, J. J., (1992). Could information theory provide an ecological theory of sensory processing? *Network*, **3**, 213-251. Doi: 10.3109/0954898X.2011.638888

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., Halgren, E., (2006). Top-down facilitation of visual recognition. *PNAS*, **103(2)**, 449-454. Doi:

<https://doi.org/10.1073/pnas.0507062103>

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8. Retrieved from <http://CRAN.R-project.org/package=lme4>

Bekhtereva, V., Pritschmann, R., Keil, A., Müller, M. M., (2018). The neural signature of extracting emotional content from rapid visual streams at multiple presentation rates: A cross-laboratory study. *Psychophysiology*, **55(12)**, e13222, doi: 10.1111/psyp.13222

Bekhtereva, V., and Müller, M. M., (2015). Affective facilitation of early visual cortex during rapid picture presentation at 6 and 15Hz. *Social, Cognitive and Affective Neuroscience*, **10(12)**, 1623-1633. Doi: 10.1093/scan/nsv058

Bekhtereva, V., and Müller, M. M., (2017). Corrigendum to: Affective facilitation of early visual cortex during rapid picture presentation at 6 and 15Hz. *Social, Cognitive and Affective Neuroscience*, **12(6)**, 1022-1023. Doi: 10.1093/scan/nsx024

Bies, A. J., Blanc-Goldhammer, D. R., Boydston, C. R., Taylor, R. P., Sereno, M. E., (2016). Aesthetic responses to exact fractals driven by physical complexity. *Frontiers in Human Neuroscience*. **10(210)**. Doi: 10.3389/fnhum.2016.00210

Brachmann, A., and Redies, C., (2017). Computational and experimental approaches to visual aesthetics. *Frontiers in Computational Neuroscience*, **11(102)**. Doi: 10.3389/fncom.2017.00102

Brady, N., and Field, D. J., (2000). Local contrast in natural images: normalisation and coding efficiency. *Perception*, **29**, 1041-1055.

Brainard, D. H. (1997). The Psychophysics Toolbox, *Spatial Vision*, **10**, 443-446. Doi: 10.1163/156856897X00357

Burton, G. J., and Moorhead, I. R., (1987). Color and spatial structure in natural scenes. *Appl Opt*, **26(1)**, 157-70. Doi: 10.1364/AO.26.000157.

Campbell, F. W., and Robson, J. G., (1968). Application of Fourier analysis for the visibility of gratings. *J. Physiol.*, **197(3)**, 551-566. Doi: 10.1113/jphysiol.1968.sp008574

Colby, E., & Bair, E. (2013). Cross-validation for nonlinear mixed effects models. *Journal of pharmacokinetics and pharmacodynamics*, **40(2)**, 243–252. <https://doi.org/10.1007/s10928-013-9313-5>

Dakin, S., Hess, R., Ledgeway, T., Achtman, R., (2002). What causes non-monotonic tuning of fMRI response to noisy images? *Current Biology*, **12 (14)**, R476–R477. Doi: 10.1016/S0960-9822(02)00960-0

Datta, R., Joshi, D., Li, J., and Wang, J. Z., (2006). Studying aesthetics in photographic images using a computational approach. In *Computer Vision, ECCV 2006. Lecture Notes in Computer Science*, **3953**, Springer, 288-301. Doi: 10.1007/11744078_23

Daugman, J. G., (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, **20**, 847-856. Doi: 10.1016/0042-6989(80)90065-6

Delorme, A., and Makeig, S., (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, **134**, 9-21. Doi: 10.1016/j.jneumeth.2003.10.009

de Rooij, M., & Weeda, W. (2020). Cross-validation: A method every psychologist should know. *Advances in Methods and Practices in Psychological Science*, **3(2)**, 248-263. Doi: <https://doi.org/10.1177/2515245919898466>

Fernandez, D., and Wilkins, A. J., (2008). Uncomfortable images in art and nature. *Perception*, **37(7)**, 1098-1113, doi: 10.1068/p5814

Field, D. J., (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A.*, **4(12)**, 2379-2394. Doi: 10.1364/JOSAA.4.002379

Field, D. J., (1993). Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mammalian visual systems. In *Wavelets, Fractals and Fourier Transforms: New Developments and New Applications*. Farge, M., Hunt, J., Vassilicos, J., (eds), Oxford Univeristy Press.

Field, D. J., (1994). What is the goal of sensory coding? *Neural Computation*, **6**, 559-601. Doi: 10.1162/neco.1994.6.4.559

Field, D. J., (1999). Wavelets, vision and the statistics of natural scenes. *Phil Trans R Soc Lond A.*, **357**, 2527-2542. Doi: 10.1098/rsta.1999.0446

Field, D. J., and Brady, N., (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, **37(23)**, 3367-3383. Doi: 10.1016/S0042-6989(97)00181-8

Graham, D. J., and Field, D. J., (2007). Statistical regularities of art images and natural scenes. Spectra, sparseness and nonlinearities. *Spatial Vision*, **21**, 149-164. Doi: 10.1163/156856807782753877

Graham, D. J., and Field, D. J., (2008). Variations in intensity statistics for representational and abstract art, and for art from the eastern and western hemispheres. *Perception*, **37**, 1341-1352. doi: 10.1068/p5971

Gratton, G., Coles, M. G., and Donchin, E., (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol*, **55(4)**, 468-84. Doi: 10.1016/0013-4694(83)90135-9

Grebenkina, M., Brachmann, A., Bertamini, M., Kadhum, A., Redies, C., (2018). Edge-orientation entropy predicts preference for diverse types of man-made images. *Frontiers in Neuroscience*, **12**, 678, doi: 10.3389/fnins.2018.00678

Hagerhall, C. M., Laike, T., Kuller, M., Marcheschi, E., Boydston, C., & Taylor, R. P. (2015). Human physiological benefits of viewing nature: EEG responses to exact and statistical fractal patterns. *Nonlinear Dynamics Psychol. Life Sci*, **19**, 1-12.

Hateren, J. H. van, and Schaaf, A. van der, (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society B*, **265(1394)**, 359-366. Doi: 10.1098/rspb.1998.0303

Hermens, F., & Zanker, J. (2012). Looking at Op Art: Gaze stability and motion illusions. *i-Perception*, **3(5)**, 282-304. Doi: 10.1068/i0457aap

Hibbard, P. B., & O'Hare, L. (2015). Uncomfortable images produce non-sparse responses in a model of primary visual cortex. *Royal Society Open Science*, **2(2)**. Doi: 10.1098/rsos.140535

Juricevic, I., Land, L., Wilkins, A., Webster, M. A., (2010). Visual discomfort and natural image statistics. *Perception*, **39(7)**, 884-99. doi: 10.1068/p6656

- Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T., ... & Berman, M. G. (2015). Is the preference of natural versus man-made scenes driven by bottom-up processing of the visual features of nature?. *Frontiers in Psychology*, **6**, 471. Doi: <https://doi.org/10.3389/fpsyg.2015.00471>
- Kass, R. E., and Raftery, A. E., (1995). "Bayes Factors", *Journal of the American Statistical Association*, **90(430)**, 773-795, doi: 10.2307/2291091
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, **36(14)**, 1-16. Doi: 10.1068/v070821
- Koch, M., Denzler, J., and Redies, C. (2010). 1/f² characteristics and isotropy in the fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PLoS ONE* **5**, e12268. Doi: 10.1371/journal.pone.0012268
- Lennie, P., (2003). The cost of cortical computation. *Current Biology*, **13(6)**, 493-7. Doi: 10.1016/S0960-9822(03)00135-0.
- Li, J., Du, Q., & Sun, C. (2009). An improved box-counting method for image fractal dimension estimation. *Pattern Recognition*, **42(11)**, 2460-2469. Doi: 10.1016/j.patcog.2009.03.001
- Locher, P. (2020). The Study of Eye Movements in Empirical Aesthetics. In: *The Oxford Handbook of Empirical Aesthetics*. Doi: 10.1093/oxfordhb/9780198824350.013.12
- McManus, C., I., Cheema, B., & Stoker, J. (1993). The Aesthetics of Composition: A Study of Mondrian. *Empirical Studies of the Arts*, **11(2)**, 83-94. Doi: 10.2190/HXR4-VU9A-P5D9-BPQQ

Mather, G. (2013). *The psychology of visual art: eye, brain and art*. Cambridge University Press, page 149

Mather, G. (2014). Artistic adjustment of image spectral slope. *Art and Perception*, **2**, 11-22.

Doi: 10.1163/22134913-00002018

Menzel, C., Hayn-Leichsenring, G. U., Langner, O., Wiese, H., and Redies, C. (2015).

Fourier power spectrum characteristics of face photographs: attractiveness perception depends on low-level image properties. *PLoS ONE* **10**, e0122801. doi:

10.1371/journal.pone.0122801

Menzel, C., Kovács, G., Amado, C., Hayn-Leichsenring, G. U., Redies, C., (2018). Visual mismatch negativity indicates automatic task-independent detection of artistic image

composition in abstract artworks. *Biological Psychology*, **136**, 76-86. doi:

10.1016/j.biopsycho.2018.05.005

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., Woods, D. L., (2002). Shape perception reduces activity in human primary visual cortex. *PNAS*, **99**, 15164-1569. Doi:

<https://doi.org/10.1073/pnas.192579399>

Nadal, M. (2013). The experience of art: insights from neuroimaging. *Progress in Brain*

Research, **204**, 135-158. Doi : <https://doi.org/10.1016/B978-0-444-63287-6.00007-5>

Nayak B.K., Karmakar S. (2019) Eye Tracking Based Objective Evaluation of Visual

Aesthetics: A Review. In: Rebelo F., Soares M. (eds) *Advances in Ergonomics in Design*.

AHFE 2018. *Advances in Intelligent Systems and Computing*, **777**. Springer, Cham.

https://doi.org/10.1007/978-3-319-94706-8_41

Norcia, A. M., Appelbaum, L., Ales, J. M., Cottureau, B. R., Rossion, B., (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, **15(4)**, doi: 10.1167/15.6.4

Ogawa, N., and Motoyoshi, I., (2020). Differential effects of orientation and spatial frequency spectra on visual unpleasantness. *Frontiers in Psychology*, **11**, 1342. Doi: <https://doi.org/10.3389/fpsyg.2020.01342>

O'Hare, L., and Hibbard, P. B., (2011). Spatial frequency and visual discomfort. *Vision Research*, **51(15)**, 1767-1777. Doi: 10.1016/j.visres.2011.06.002

O'Hare, L., and Hibbard, P. B., (2013). Visual discomfort and blur. *Journal of Vision*, **13(5)**, doi 10.1167/13.5.7.

O'Hare, L., Hibbard, P. B., Zhang, T. & Nefs, H. T., (2013). Visual discomfort and depth-of-field. *i-Perception*, **4(3)**, 156–169. Doi: 10.1068/i0566

O'Hare, L., and Goodwin, P., (2018). ERP responses to images of abstract artworks, photographs of natural scenes, and artificially created uncomfortable images. *Journal of Cognitive Psychology*, **30(5-6)**, 627-641, doi: 10.1080/20445911.2018.1499657.

Olmos, A., and Kingdom, F. A. A., (2004). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, **33**, 1463-1473. Doi: 10.1068/p5321

Olshausen, B. A., and Field, D. J., (2000). Vision and the coding of natural images. *American Scientist*, **88**, 238-245. Doi: 10.1511/2000.3.238

Olshausen, B. A., and Field, D. J., (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, **14(4)**, 481-7. Doi: 10.1016/j.conb.2004.07.007

Oppenheim A. V. Lim J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, **69 (5)**, 529–541. Doi: 10.1109/PROC.1981.12022

Párraga, C. A., Troscianko, T., & Tolhurst, D. J. (2000). The human visual system is optimised for processing the spatial information in natural visual images. *Current Biology*, **10(1)**, 35-38. Doi: [https://doi.org/10.1016/S0960-9822\(99\)00262-6](https://doi.org/10.1016/S0960-9822(99)00262-6)

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision* **10**, 437-442. Doi: 10.1163/156856897X00366

Penacchio, O., and Wilkins, A. J., (2015). Visual discomfort and the spatial distribution of Fourier energy. *Vision Research*, **108**, 1-7. Doi: 10.1016/j.visres.2014.12.013

Pentland, A. P., (1984). Fractal-based description of natural scenes. *IEE Transactions on Pattern Analysis and Machine Intelligence*, **6(6)**, 661-674. Doi: 10.1109/TPAMI.1984.4767591

Plant, G. T., (1983). Transient visually evoked potentials to sinusoidal gratings in optic neuritis. *Journal of Neurology, Neurosurgery and Psychiatry*, **46**, 1125-1133. Doi: 10.1136/jnmp.46.12.1125

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available from: <http://www.R-project.org/>.

Ramachandran, V. S., and Hirstein, W., (1999). The science of art: A neurological theory of aesthetic experience. *Journal of Consciousness Studies*. **6(6-7)**, 15-41. Doi: <https://www.ingentaconnect.com/content/imp/jcs/1999/00000006/f0020006/949>

Redies, C., (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision*, **21(1-2)**, 97-117. Doi: 10.1068/p5814

Redies, Hasenstein, J., and Denzler, J., (2007). Fractal-like image statistics in visual art: similarity to natural scenes. *Spatial Vision* **21(1-2)**, 137-148. Doi: 10.1163/156856807782753921

Redies, C., (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Frontiers in Human Neuroscience*, **9(218)**, doi:10.3389/fnhum.2015.00218

Ruderman, D. L., (1994). The statistics of natural images. *Network: Comput Neural Systems*, **5**, 517-548. Doi: https://doi.org/10.1088/0954-898X_5_4_006

Sarasso, P., Ronga, I., Kobau, P., Bosso, T., Artusio, I., Ricci, R., Neppi-Modona, M., (2020). Beauty in mind: Aesthetic appreciation correlates with perceptual facilitation and attentional amplification. *Neuropsychologia*, **136**, 107282. Doi: <https://doi.org/10.1016/j.neuropsychologia.2019.107282>

Shamsyeh Zahedi, M., & Zeil, J. (2018). Fractal dimension and the navigational information provided by natural scenes. *PLoS ONE*, **13(5)**, e0196227. Doi: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0196227>

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, **24(1)**, 1193-1216. Doi: <https://doi.org/10.1146/annurev.neuro.24.1.1193>

Singmann, H., Bolker, B., Westfall, J., and Aust, F., (2016). afex: Analysis of Factorial Experiments. R package version 0.16-1. <https://CRAN.R-project.org/package=afex>

Song, Q. C., Tang, C., Wee, S., (2021). Making sense of model generalizability: A tutorial on cross-validation in R and Shiny. *Advances in Methods and Practices in Psychological Science*, Doi: <https://doi.org/10.1177/2515245920947067>

Spehar, B., Clifford, C. W. G., Newell, B. R., Taylor, R. P., (2003). Universal aesthetics of fractals. *Computers and Graphics*, **27(5)**, 813-820. doi: 10.1016/S0097-8493(03)00154-7

Spehar, B., Wong, S., van de Klundert, S., Lui, J., Clifford, C. W. G., Taylor, R. P., (2015). Beauty and the beholder: the role of visual sensitivity in visual preference. *Frontiers in Human Neuroscience*, **9(514)**. Doi: 10.3389/fnhum.2015.00514

Taylor, R. P., Micolich, A. P., and Jonas, D., (1999). Fractal analysis of Pollock's drip paintings. *Nature*, **399**, 422. Doi: 10.1038/20833

Thomson M. G. (1999). Visual coding and the phase structure of natural scenes. *Network*, **10(2)**, 123–132. Doi: 10.1088/0954-898X_10_2_302

Tolhurst, D. J., Tadmor, Y., and Chao, T., (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, **12(2)**, 229-32. Doi: 10.1111/j.1475-1313.1992.tb00296.x

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in the Enigma illusion. *PNAS*, **105**, 16033–16038. Doi: 10.1073/pnas.0709389105

Vartanian, O., & Goel, V. (2004). Neuroanatomical correlates of aesthetic preference for paintings. *Neuroreport*, **15(5)**, 893-897. Doi : 10.1097/00001756-200404090-00032

Vialette, F. B., Maurice, M., Dauwels, J., Cichocki, A., (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog. Neurobiol.*, **90(4)**, 418-38. doi: 10.1016/j.pneurobio.2009.11.005.

Wade, N. J., (1978). Op art and visual perception. *Perception*, **7**, 21-46. Doi: 10.1068/p070021

Wilkins, A., Nimmo-Smith, I. A. N., Tait, A., McManus, C., Della Sala, S., Tilley, A., Arnold, K., Barrie, M., Scott, S., (1984). A neurological basis for visual discomfort. *Brain*, **107(4)**, 989-1017. Doi: 10.1093/brain/107.4.989

Wilkins, A. J., Smith, J., Willison, C. K., Beare, T., Boyd, A., Hardy, G., Mell, L., Peach, C., Harper, S., (2007). Stripes within words affect reading. *Perception*, **36(12)**, 1788-1803. Doi: <https://doi.org/10.1068/p5651>

Winter, B., (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv: 1308.5499. [<http://arxiv.org/pdf/1308.5499.pdf>]

Wu, Z., (2016). Physical connections between different SSVEP neural networks. *Scientific Reports*, **6**, 22801, doi: 10.1038/srep22801.

Yanulevskaya, V., Uijlings, J., Bruni, E., Sartori, A., Zamboni, E., Bacci, F., ... & Sebe, N. (2012, October). In the eye of the beholder: employing statistical analysis and eye tracking for analyzing abstract paintings. *In Proceedings of the 20th ACM international conference on Multimedia* (pp. 349-358). ACM.

Yoshimoto, S., Garcia, J., Jiang, F., Wilkins, A. J., Takeuchi, T., Webster, M. A., (2017). Visual discomfort and flicker. *Vision Research*, **138**, 18-28. Doi: 10.1016/j.visres.2017.05.015

Zanker, J. M. (2004). Looking at op-art from a computational viewpoint. *Spatial Vision*, **17**, 75–94. Doi: 10.1163/156856804322778279

Zanker, J. M., Doyle, M., & Walker, R. (2003). Gaze stability of observers watching op art pictures. *Perception*, **32**, 1037–1049. Doi: 10.1068/p5128

Zanker, J. M., Hermens, F., & Walker, R. (2010). Quantifying and modelling the strength of motion illusions perceived in static patterns. *Journal of Vision*, **10**, 1–4. Doi: 10.1167/10.2.13



Figure 1 shows some illustrative examples of the stimuli. Top row shows artworks, middle row shows the bump stimuli, bottom row some examples of the natural images, taken from the van Hateren image database (van Hateren and van der Schaaf, 1998).

130x130mm (300 x 300 DPI)

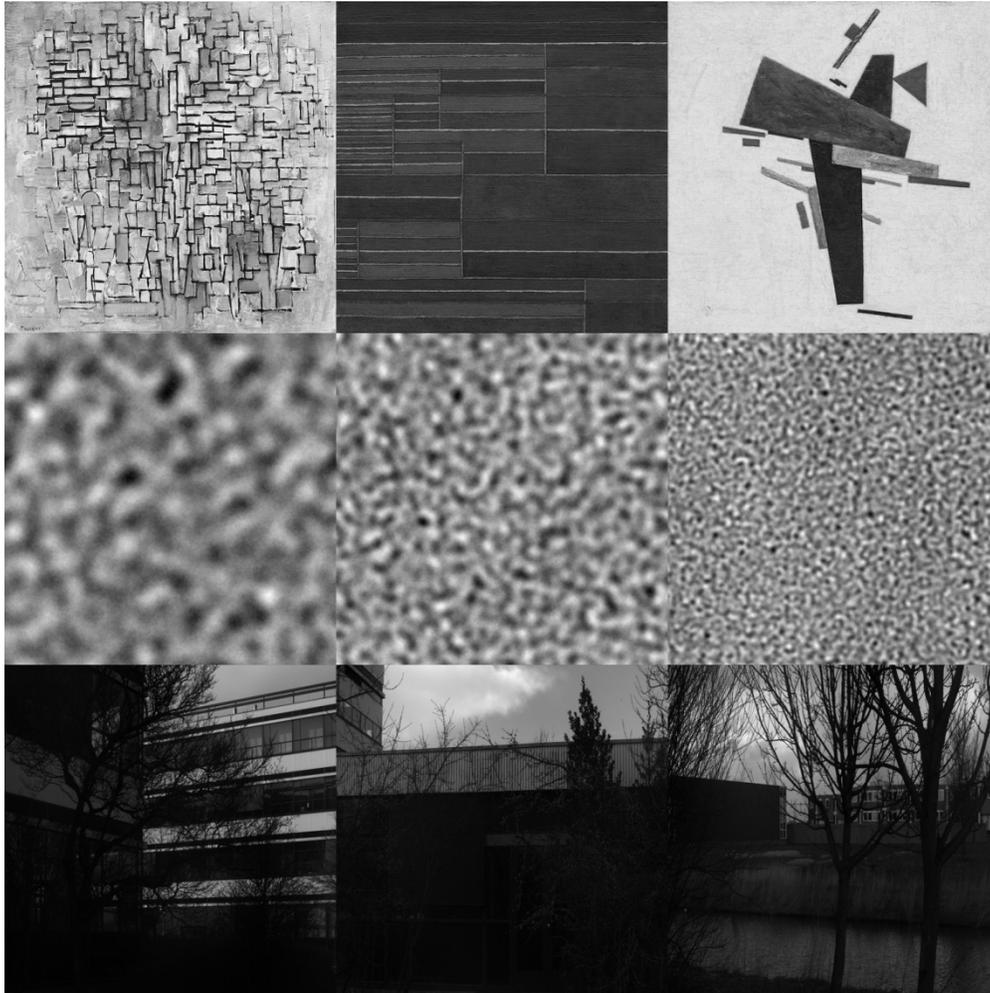


Figure 1 shows some illustrative examples of the stimuli. Top row shows artworks, middle row shows the bump stimuli, bottom row some examples of the natural images, taken from the van Hateren image database (van Hateren and van der Schaaf, 1998).

130x130mm (300 x 300 DPI)

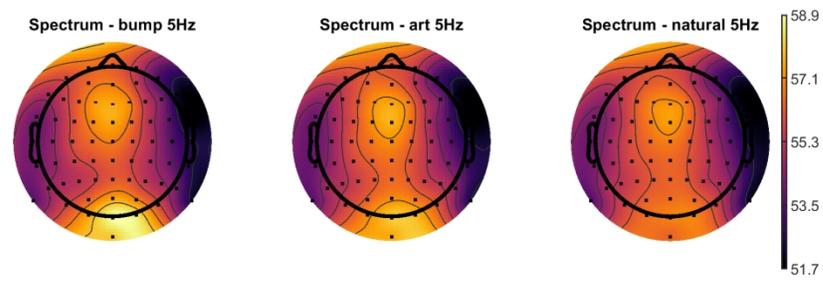


Figure 2 shows the scalp topography of SSVEP activity for the three image categories (art, bump, natural images) in Experiment 1. Scale is in $10 \cdot \log_{10}(\mu V^2/Hz)$.

177x55mm (300 x 300 DPI)

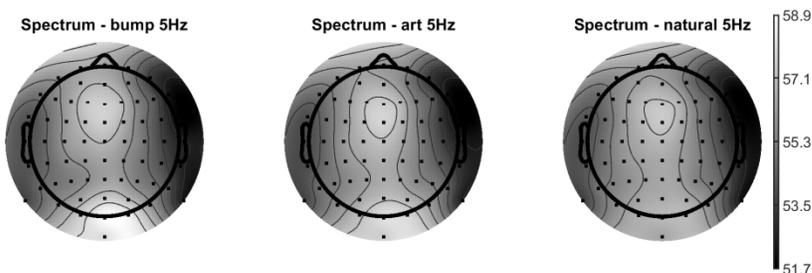


Figure 2 shows the scalp topography of SSVEP activity for the three image categories (art, bump, natural images) in Experiment 1. Scale is in $10 \cdot \log_{10}(\mu V^2/Hz)$.

177x55mm (300 x 300 DPI)

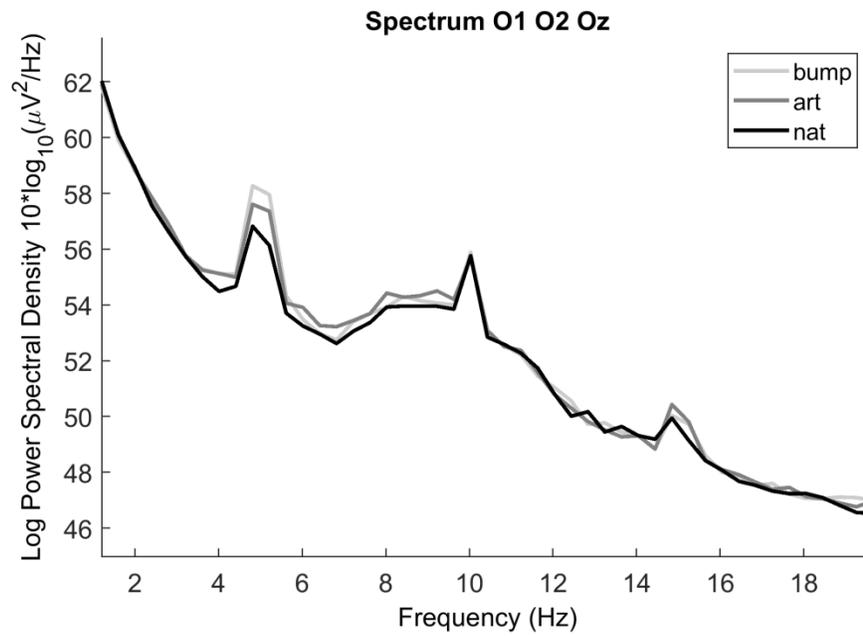


Figure 3 shows log spectral power density of the three image categories (artworks, bump stimuli and natural images). There is a clear SSVEP peak at the fundamental frequency (5Hz) and also the first and second harmonics (10 and 15Hz).

188x122mm (300 x 300 DPI)

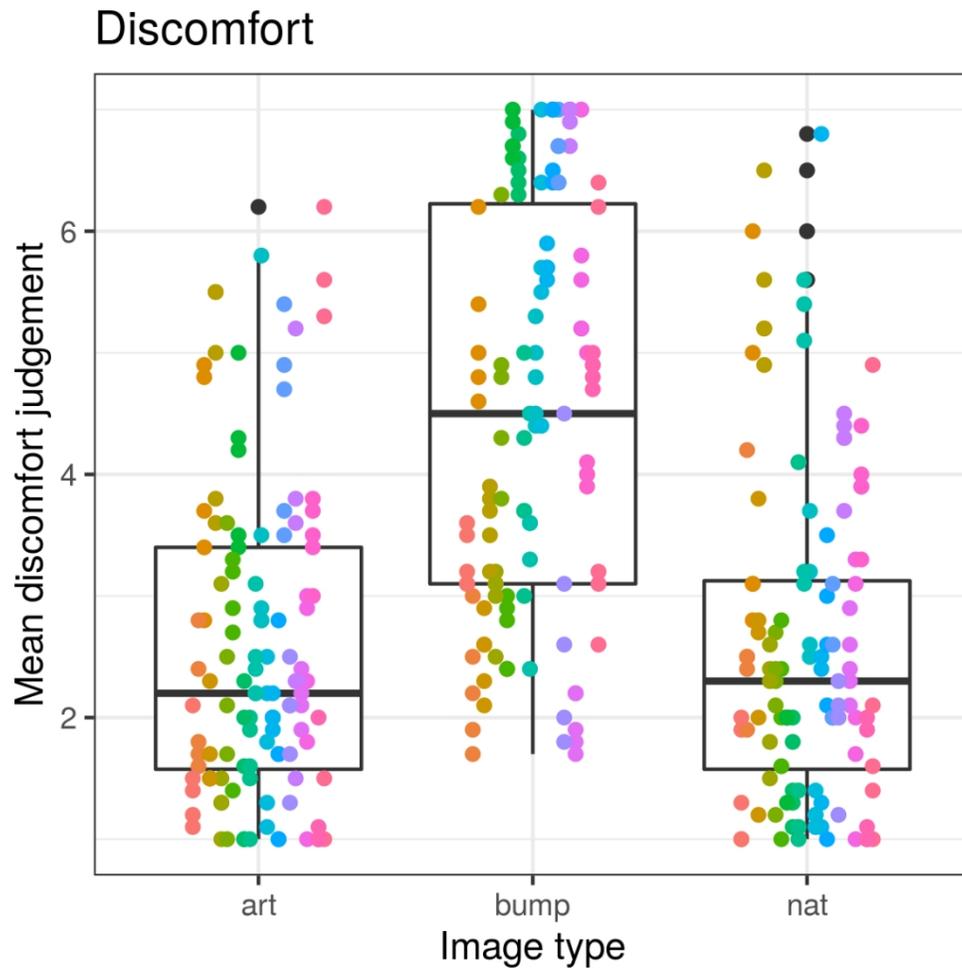


Figure 4 shows boxplots of the discomfort judgements (the mean taken from each individual) for the three image categories in Experiment 1. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

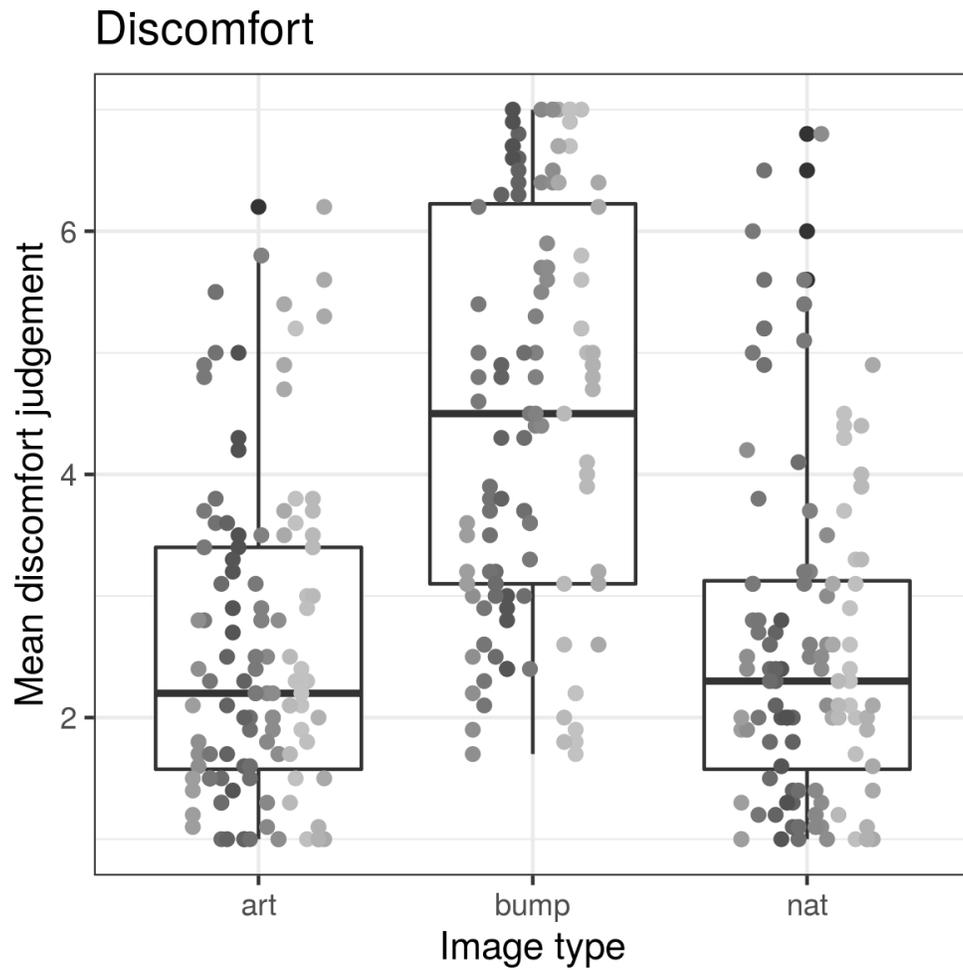


Figure 4 shows boxplots of the discomfort judgements (the mean taken from each individual) for the three image categories in Experiment 1. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

Aesthetics

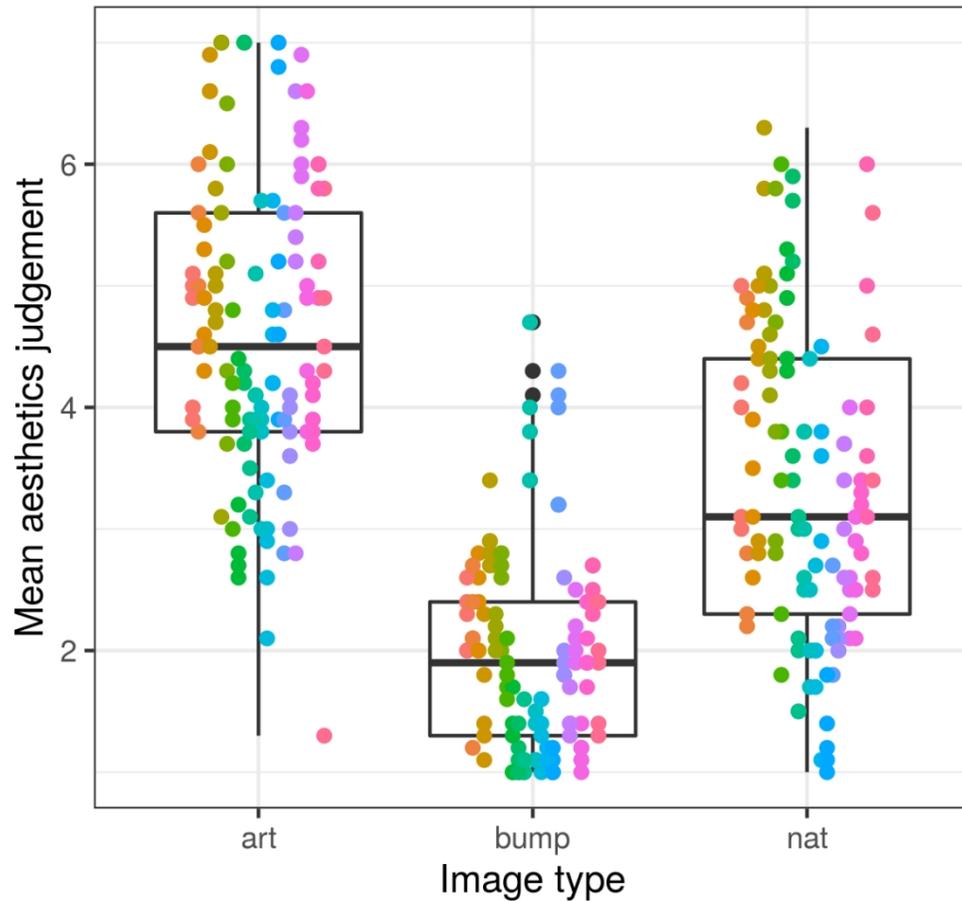


Figure 5 shows the aesthetics judgements (the mean taken from each individual) for the three image categories in Experiment 1. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

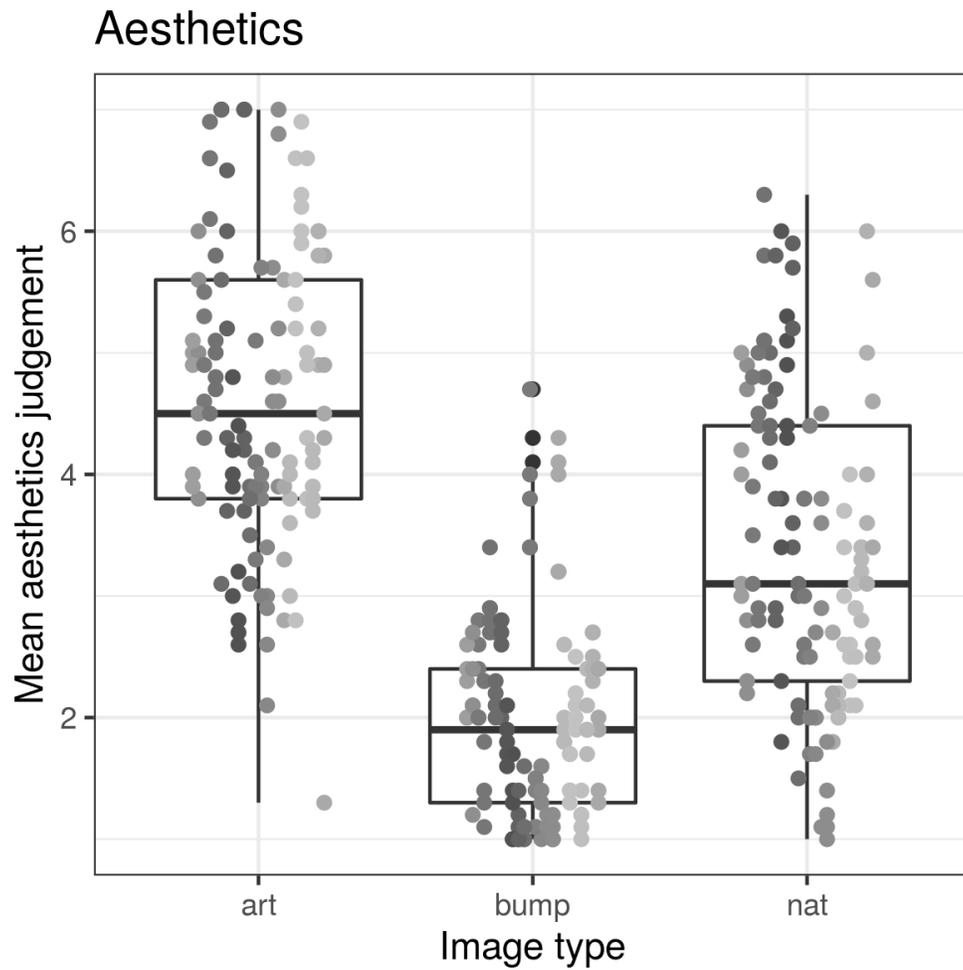


Figure 5 shows the aesthetics judgements (the mean taken from each individual) for the three image categories in Experiment 1. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

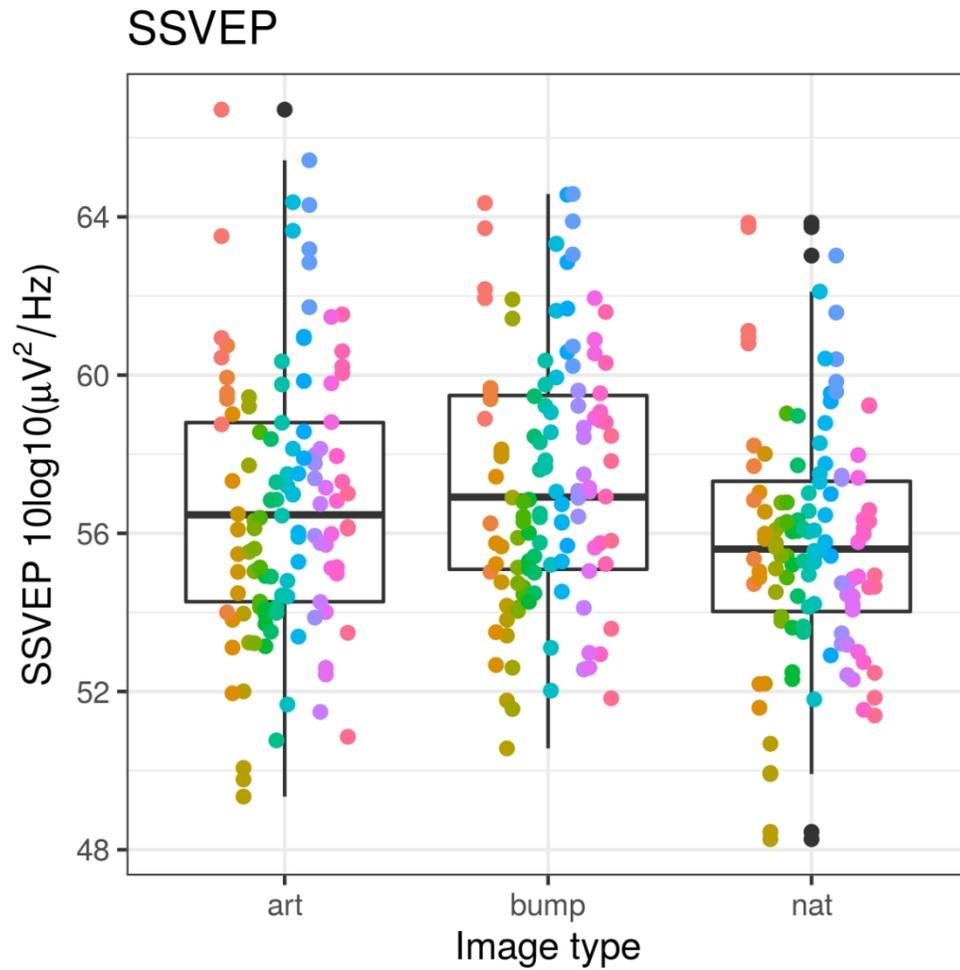


Figure 6 shows the SSVEP response (the mean for each individual) to the three image categories in Experiment 1. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

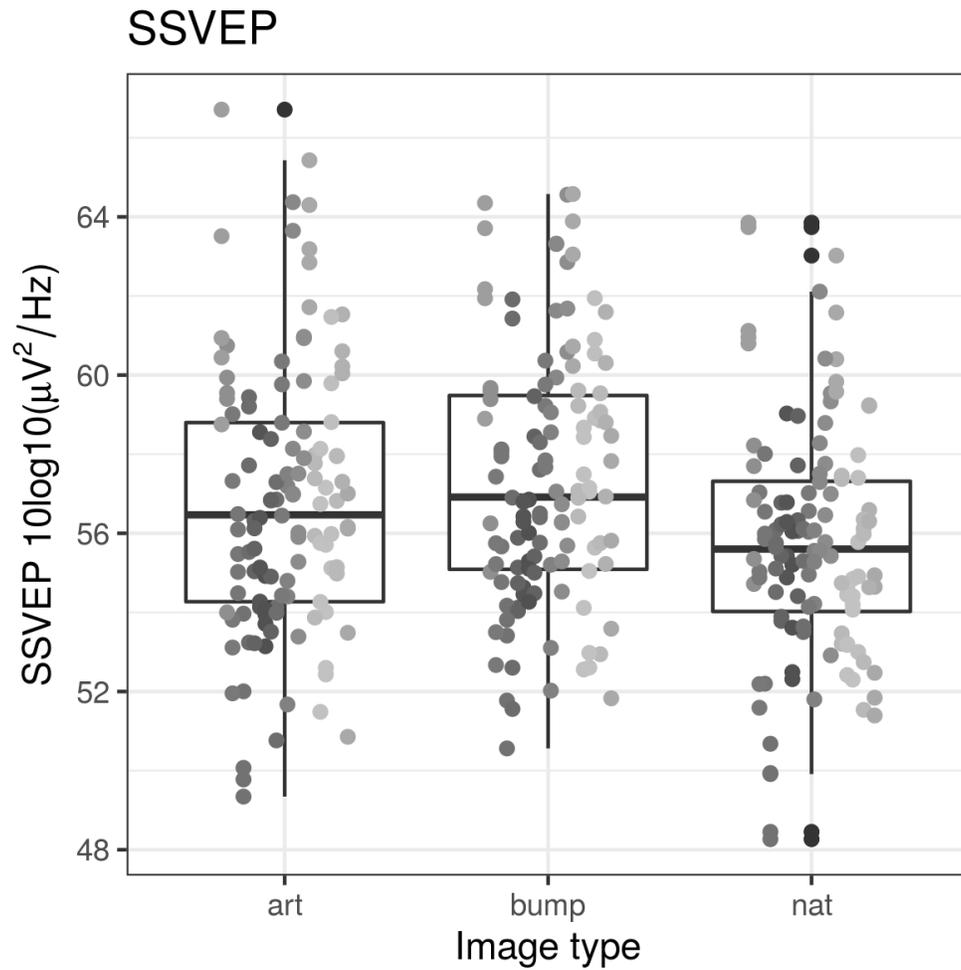


Figure 6 shows the SSVEP response (the mean for each individual) to the three image categories in Experiment 1. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

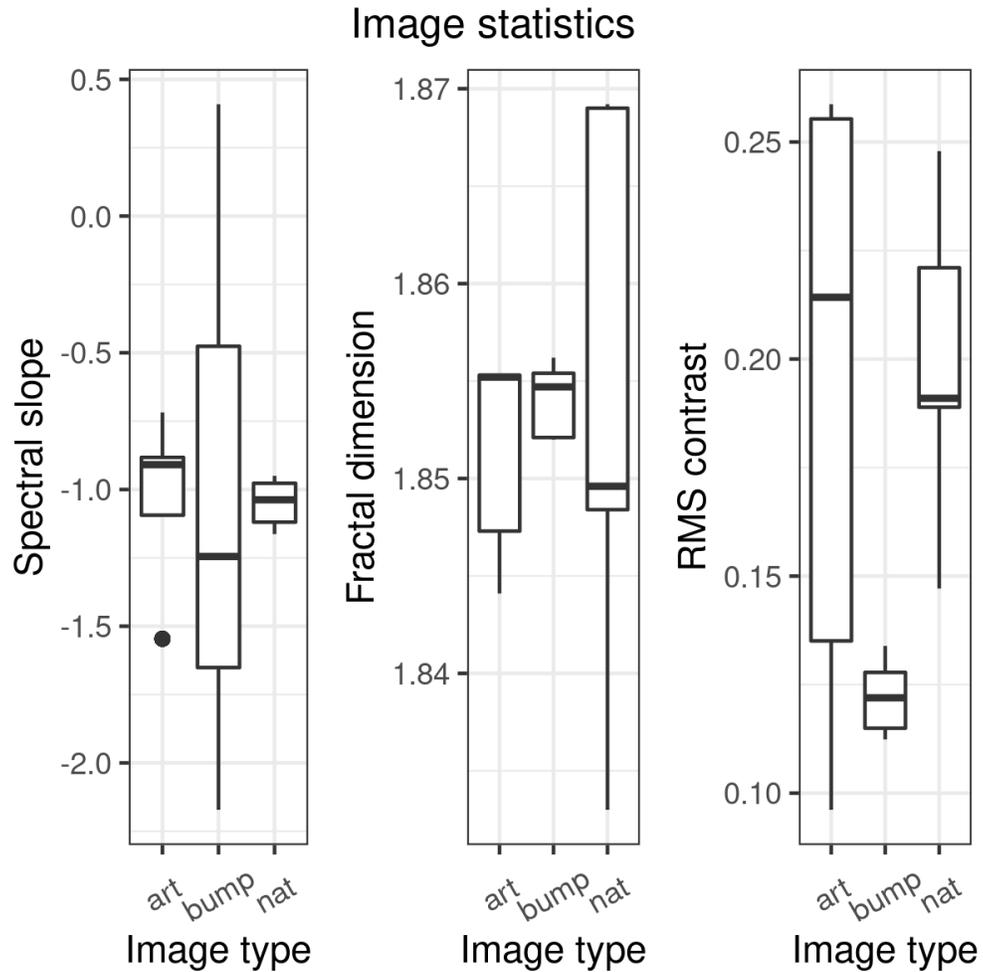


Figure 7 shows boxplots of the image statistics for the three image types. Left shows spectral slope, centre shows fractal dimension, and right shows RMS contrast. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points.

101x101mm (300 x 300 DPI)

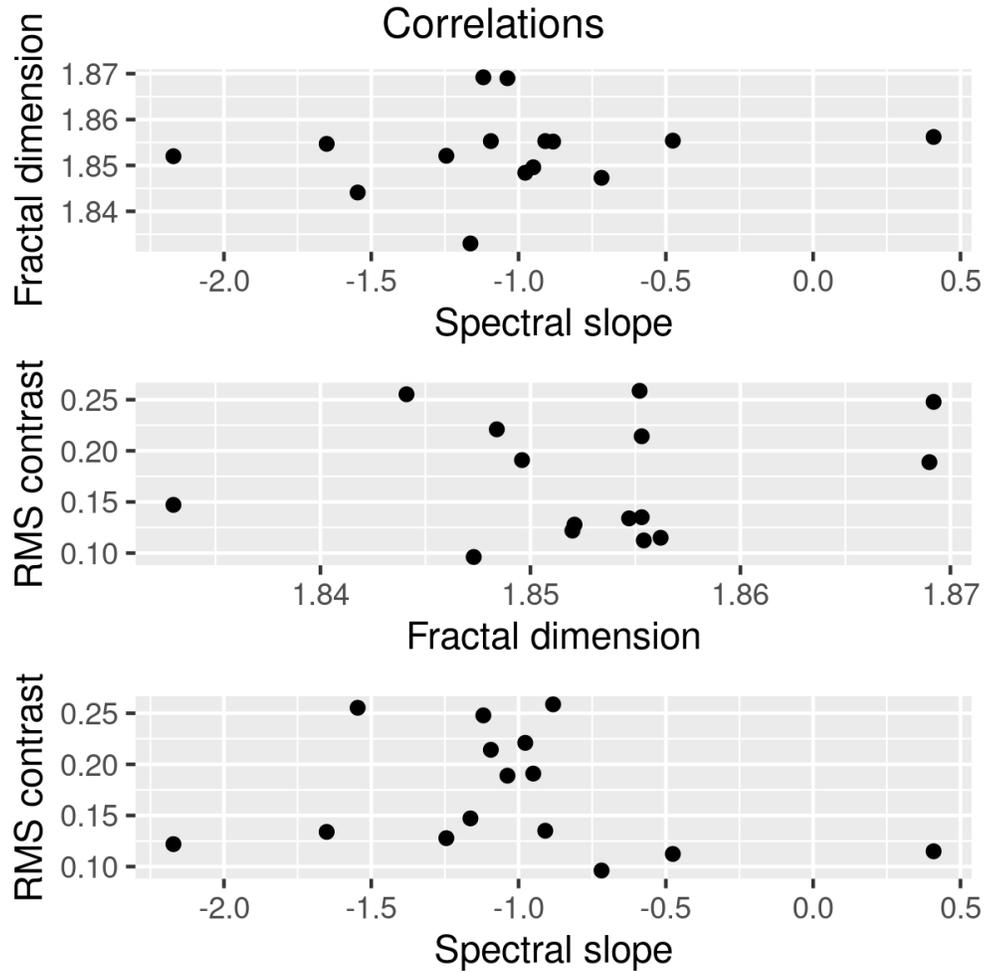


Figure 8 shows the relationship between (top) spectral slope and fractal dimension, (middle) fractal dimension and RMS contrast, and (bottom) spectral slope and RMS contrast for the images used in Experiment 1.

101x101mm (300 x 300 DPI)

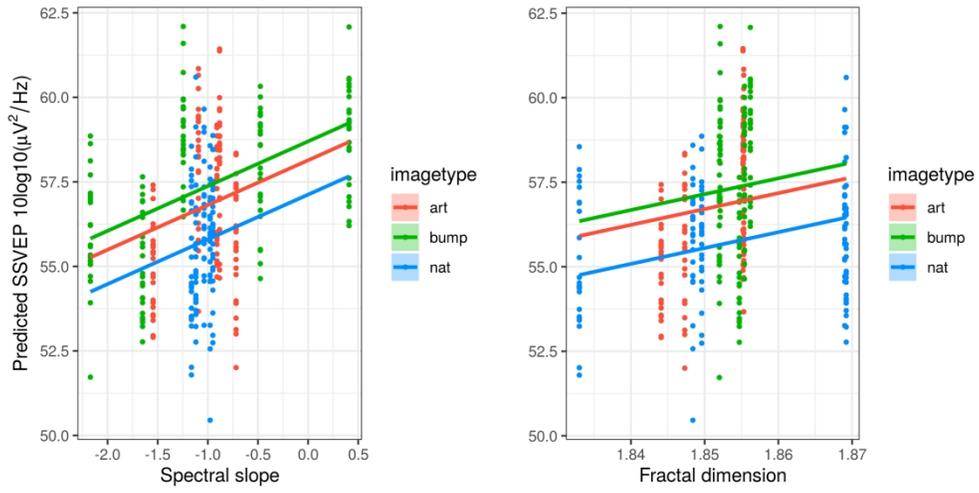


Figure 9 shows model predictions of SSVEP responses from spectral slope (left) and fractal dimension (right), including image type as a fixed effect and observer as a random effect.

203x101mm (300 x 300 DPI)

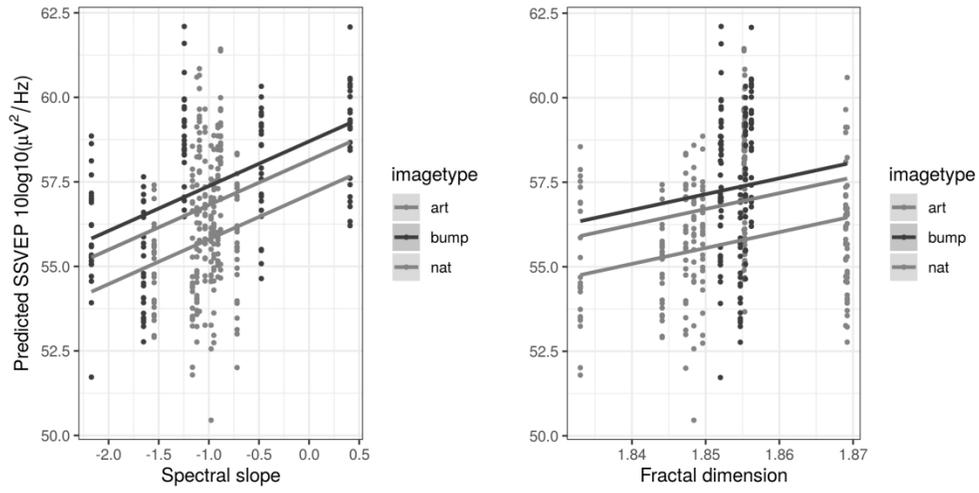


Figure 9 shows model predictions of SSVEP responses from spectral slope (left) and fractal dimension (right), including image type as a fixed effect and observer as a random effect.

203x101mm (300 x 300 DPI)

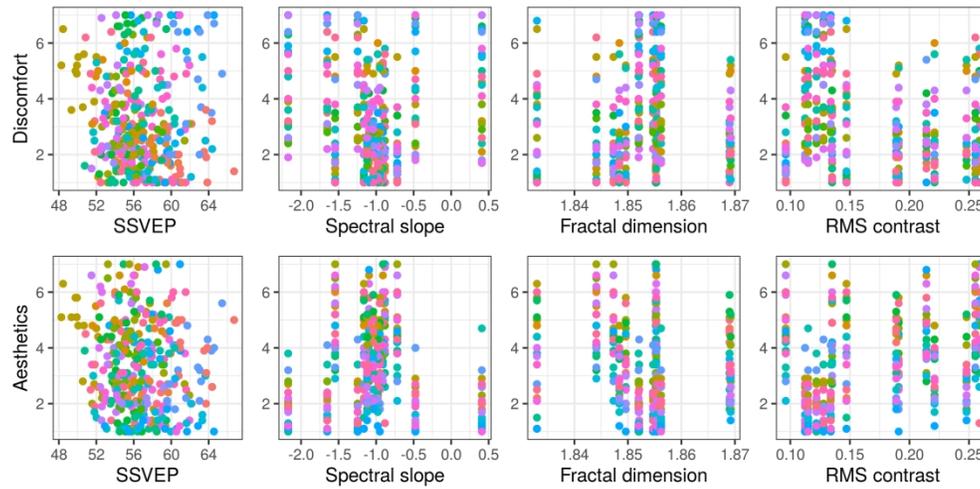


Figure 10 shows bivariate correlations of model predictors with the subjective judgements. Top row shows the model predictors for discomfort, (left to right) SSVEP, spectral slope, fractal dimension and RMS contrast. Bottom row shows model predictors for aesthetics judgements: (left to right) SSVEP, spectral slope, fractal dimension and RMS contrast. Individual observers are indicated as separate shades.

203x101mm (300 x 300 DPI)

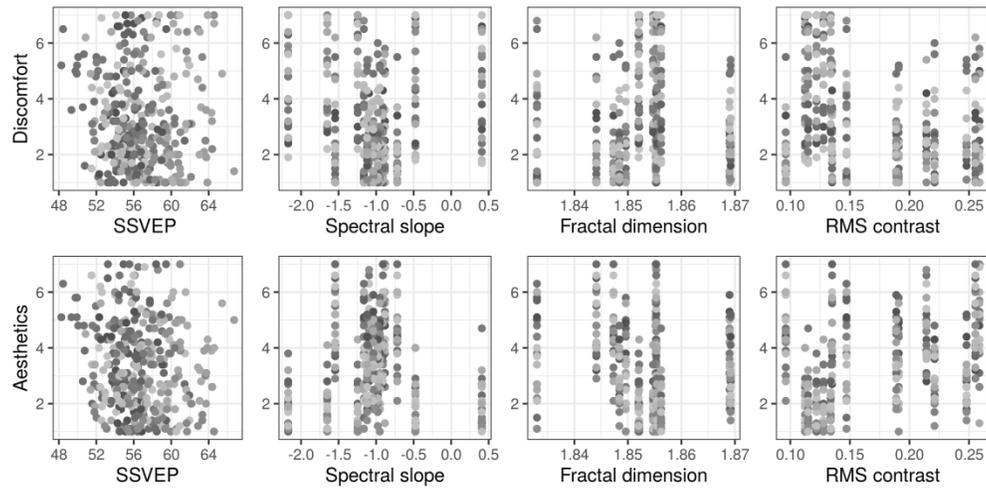


Figure 10 shows bivariate correlations of model predictors with the subjective judgements. Top row shows the model predictors for discomfort, (left to right) SSVEP, spectral slope, fractal dimension and RMS contrast. Bottom row shows model predictors for aesthetics judgements: (left to right) SSVEP, spectral slope, fractal dimension and RMS contrast. Individual observers are indicated as separate shades.

203x101mm (300 x 300 DPI)

Pleasing

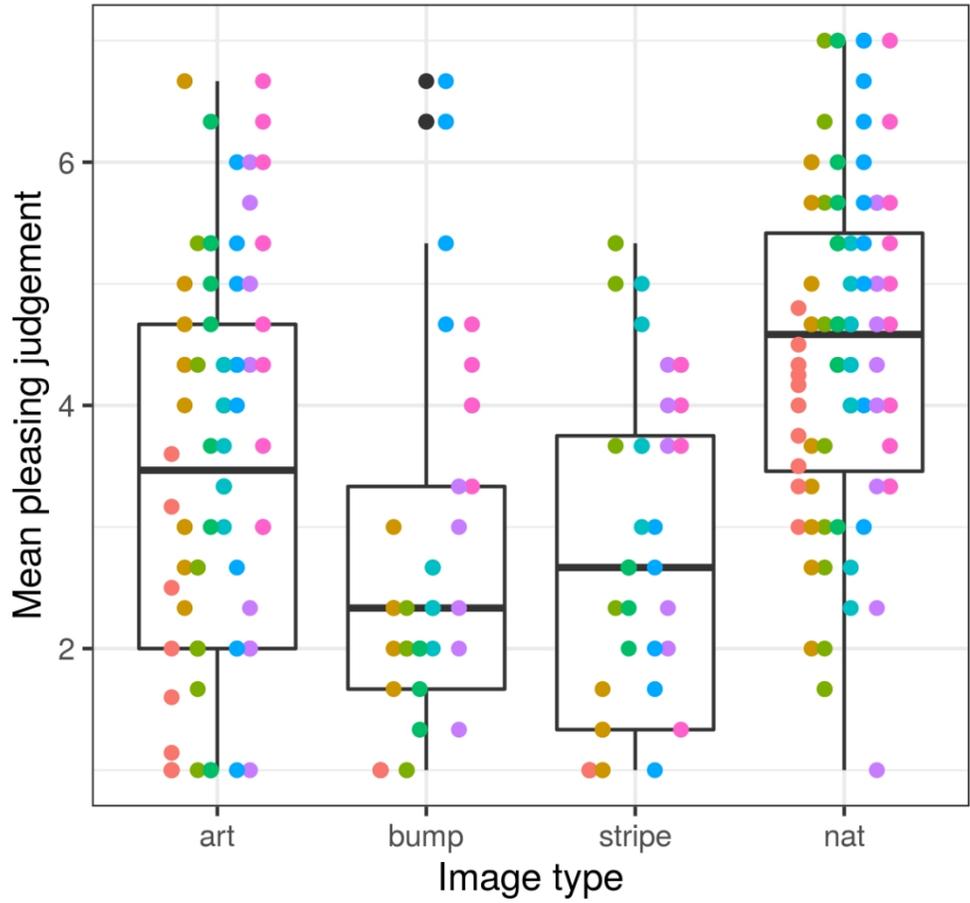


Figure 11 shows the judgements of how pleasing an image was thought to be (the mean from each individual) in Experiment 2. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

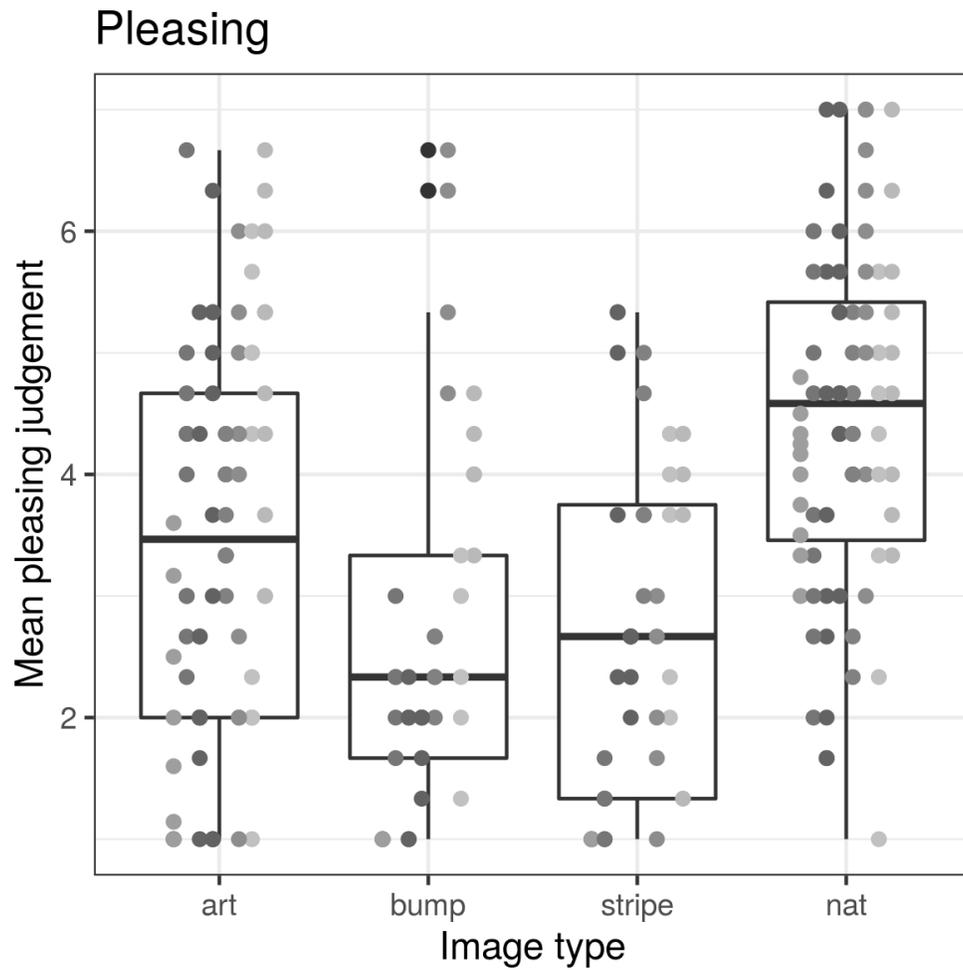


Figure 11 shows the judgements of how pleasing an image was thought to be (the mean from each individual) in Experiment 2. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

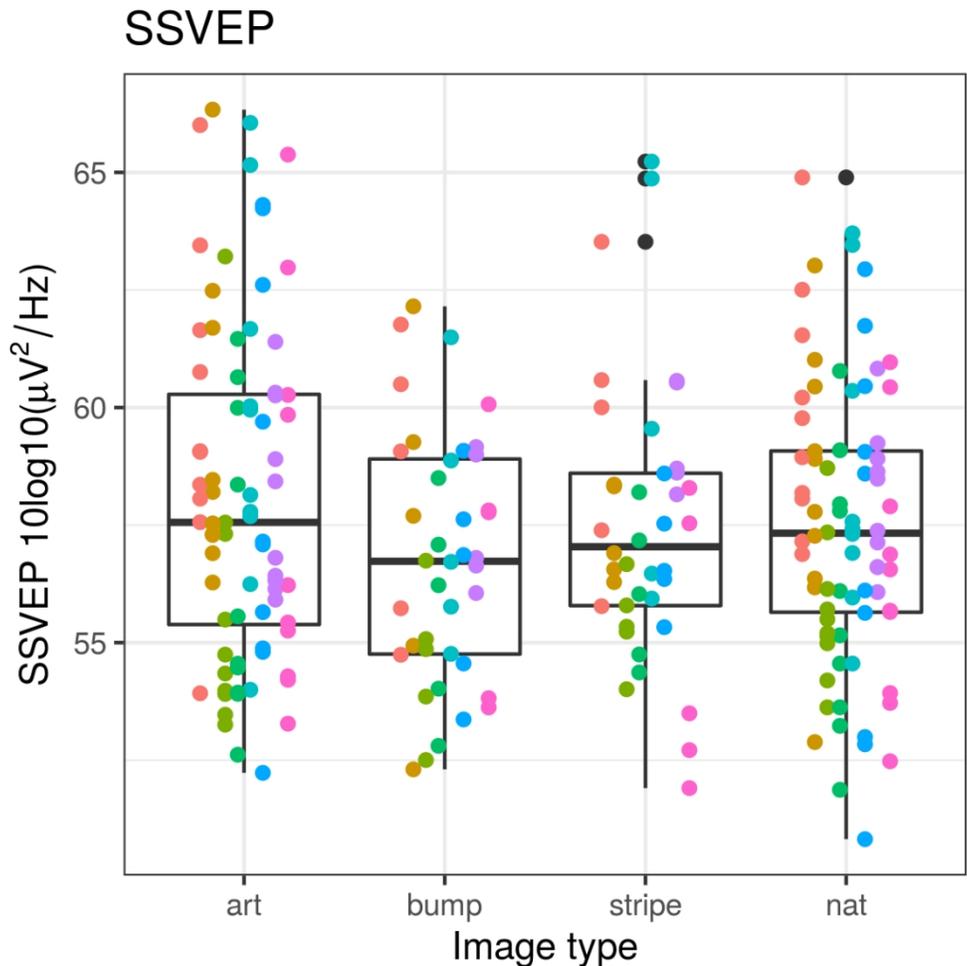


Figure 12 shows the SSVEP response (averaged over each individual) to the four image categories in Experiment 2. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

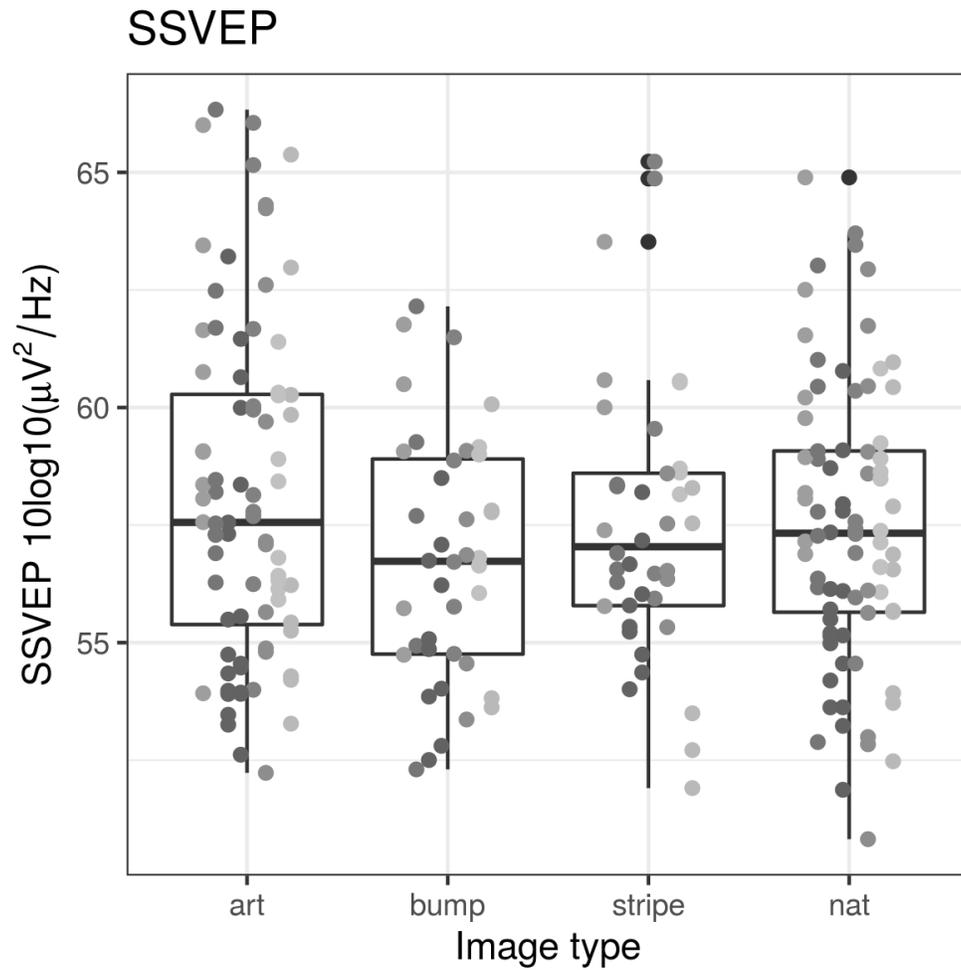


Figure 12 shows the SSVEP response (averaged over each individual) to the four image categories in Experiment 2. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

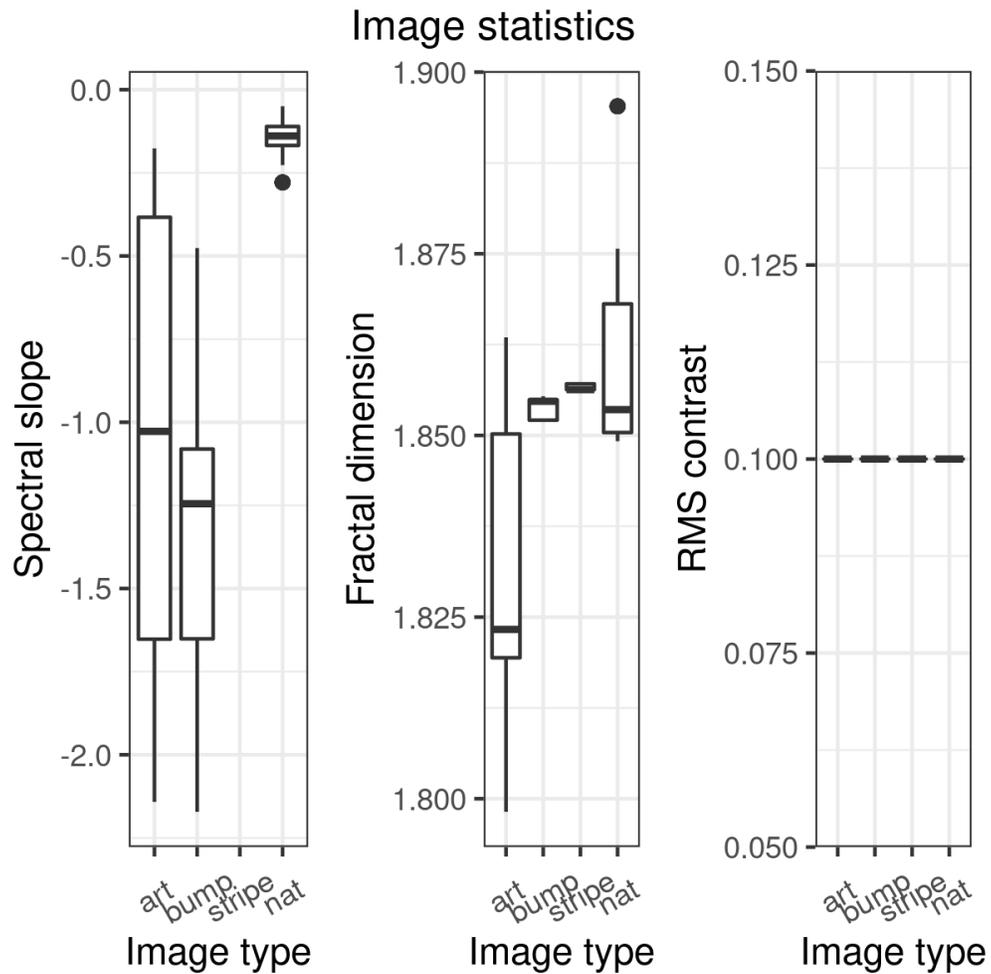


Figure 13 shows boxplots of the image statistics for the four image types. Left shows spectral slope, centre shows fractal dimension, and right shows RMS contrast. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points. RMS contrast was held constant in this experiment.

101x101mm (300 x 300 DPI)

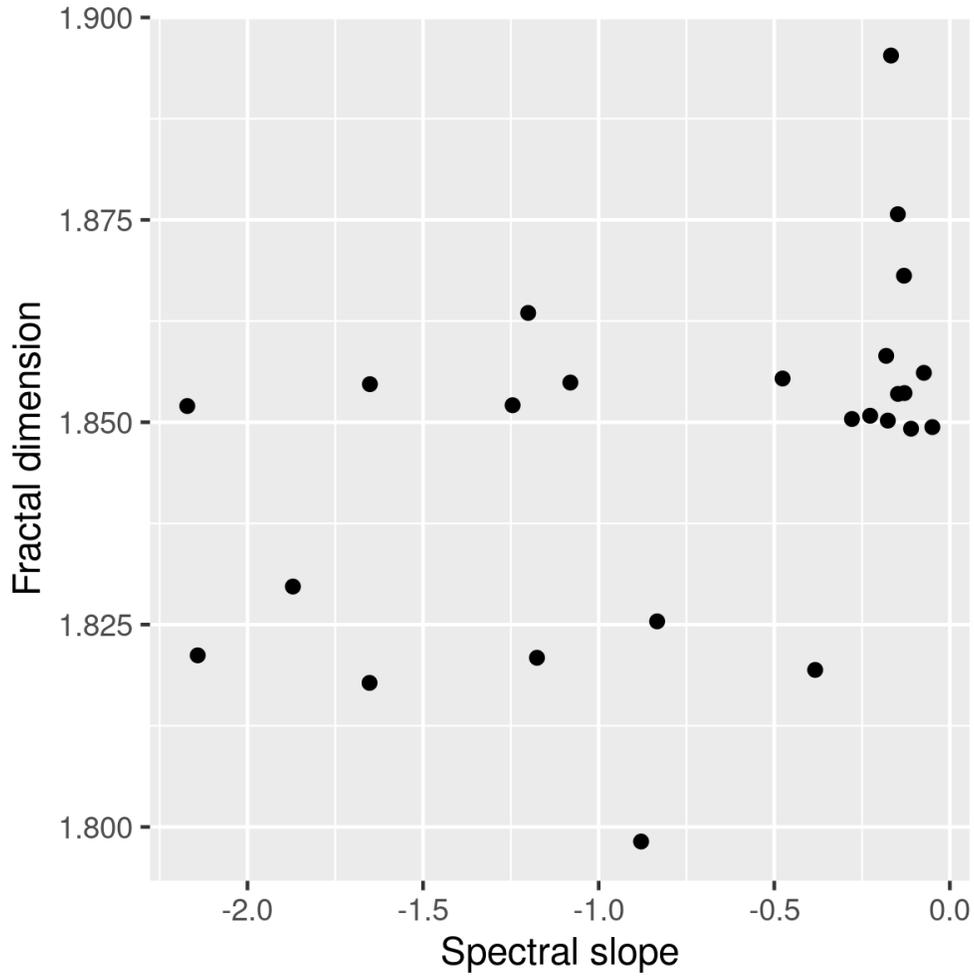


Figure 14 shows the relationship between spectral slope and fractal dimension for the images in Experiment 2.

101x101mm (300 x 300 DPI)

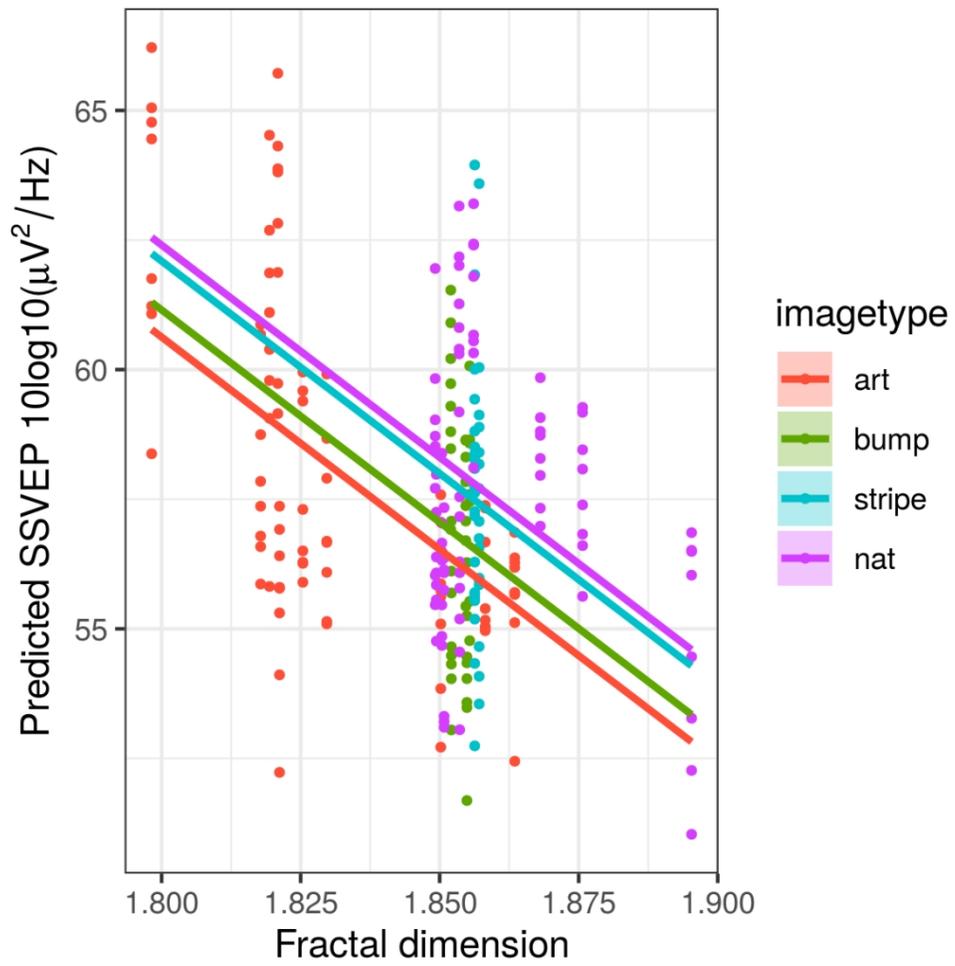


Figure 15 shows the model predicted SSVEP from fractal dimension. Image type was included as a fixed effect, and observer as a random effect.

101x101mm (300 x 300 DPI)

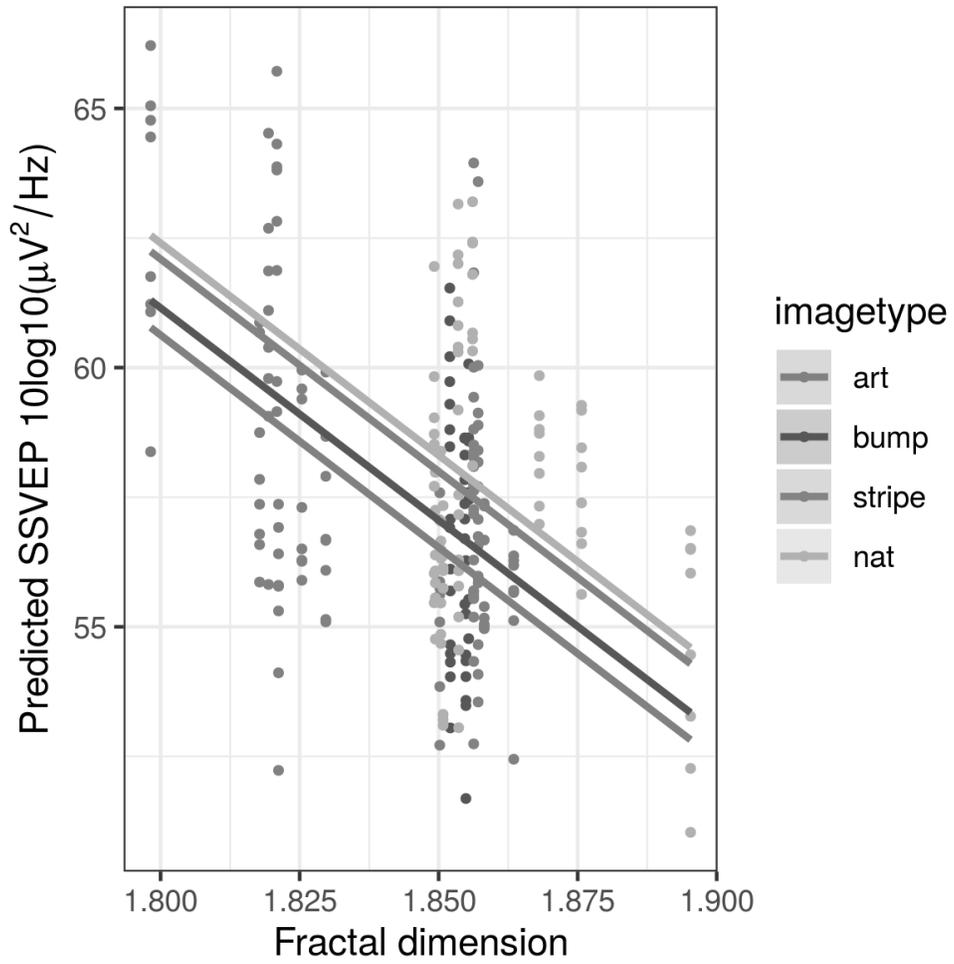


Figure 15 shows the model predicted SSVEP from fractal dimension. Image type was included as a fixed effect, and observer as a random effect.

101x101mm (300 x 300 DPI)

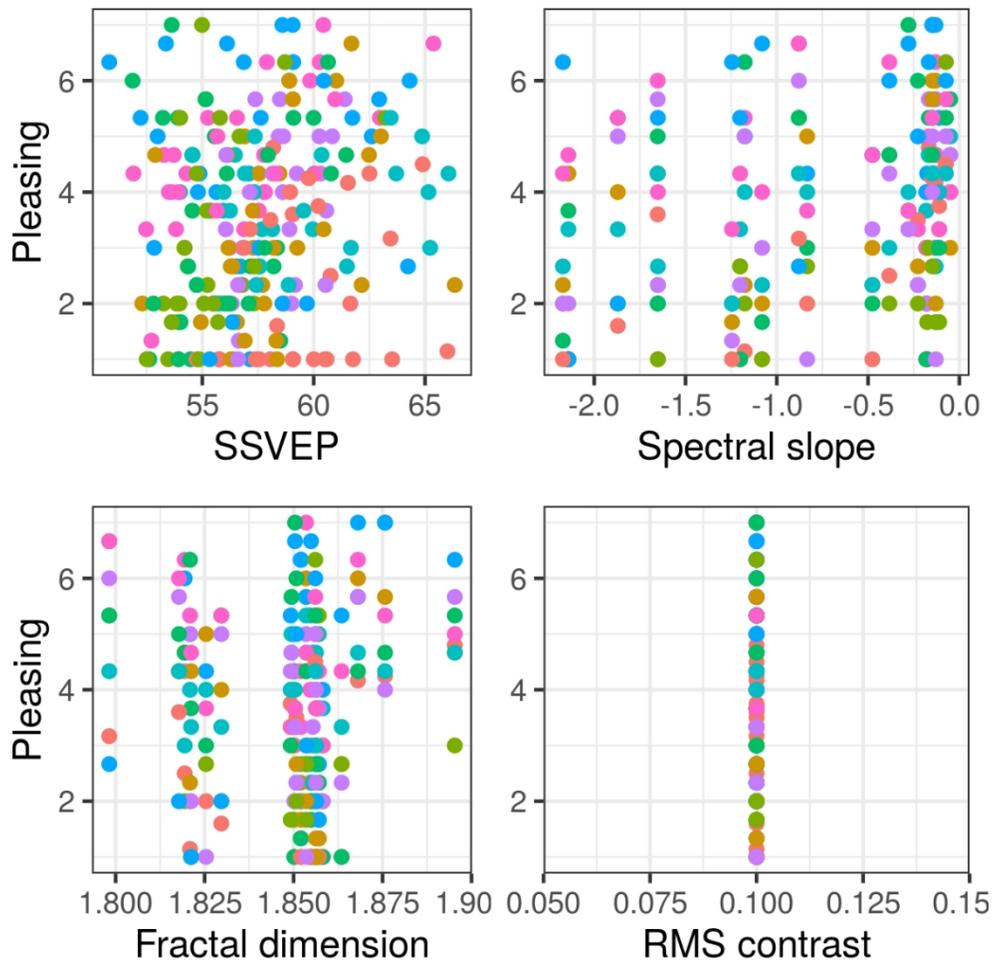


Figure 16 shows the bivariate correlations for each of the predictors included in the model against pleasing judgements. Top row shows SSVEP and spectral slope, bottom row shows fractal dimension and RMS contrast. RMS contrast was constant in this experiment. Each individual is represented with a different shade.

101x101mm (300 x 300 DPI)

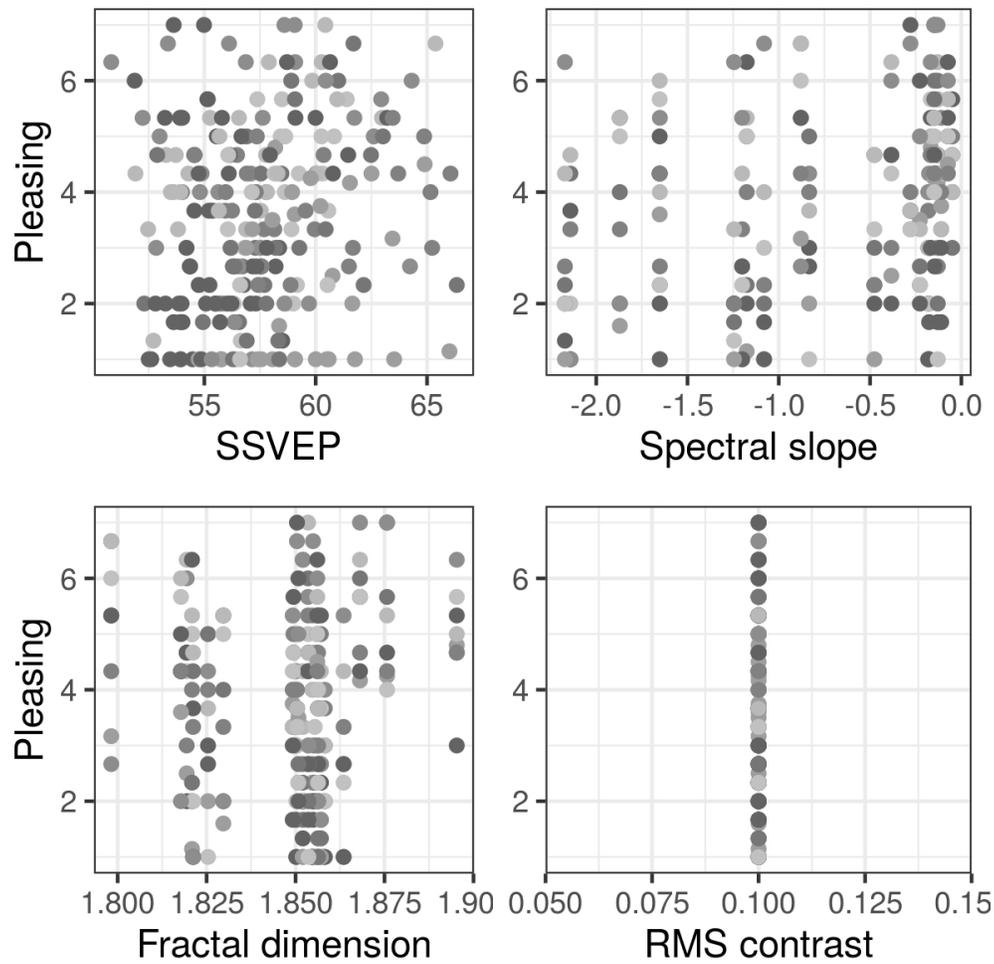


Figure 16 shows the bivariate correlations for each of the predictors included in the model against pleasing judgements. Top row shows SSVEP and spectral slope, bottom row shows fractal dimension and RMS contrast. RMS contrast was constant in this experiment. Each individual is represented with a different shade.

101x101mm (300 x 300 DPI)

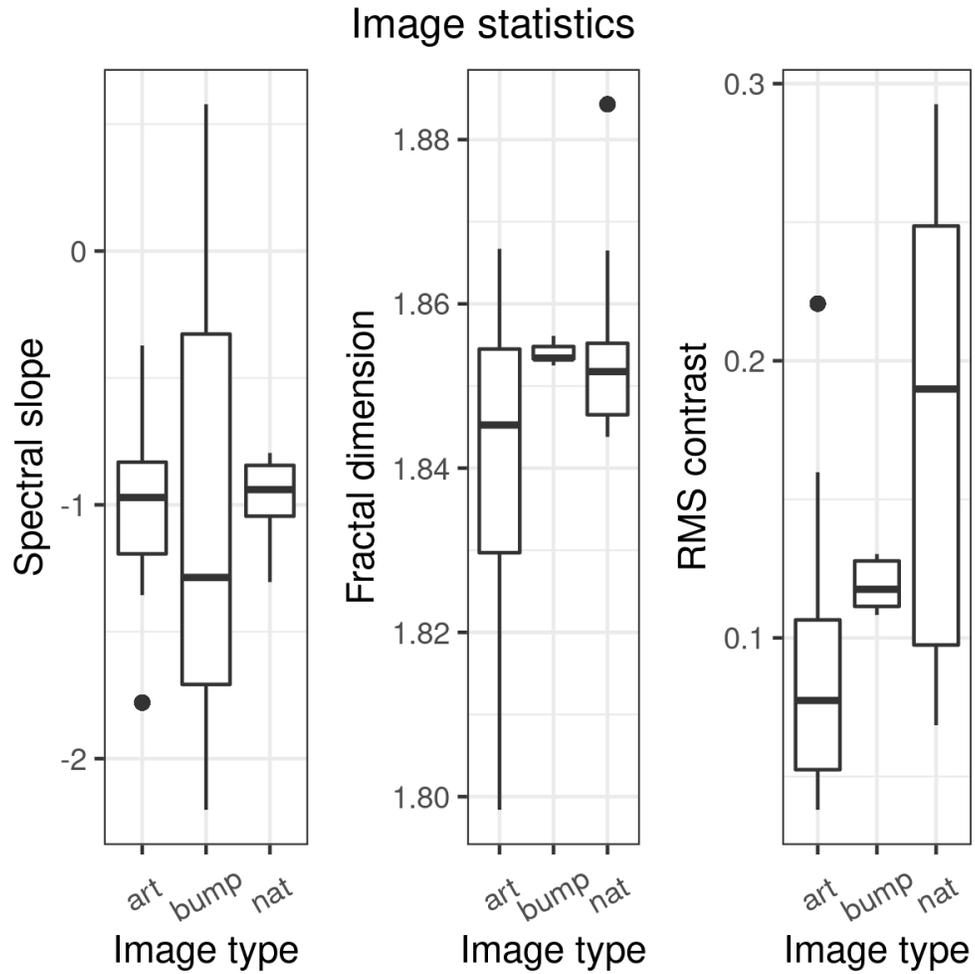


Figure 17 shows box plots of the image statistics for the three image types used in Experiment 3. Left shows spectral slope, middle shows fractal dimension, and right shows RMS contrast. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points.

101x101mm (300 x 300 DPI)

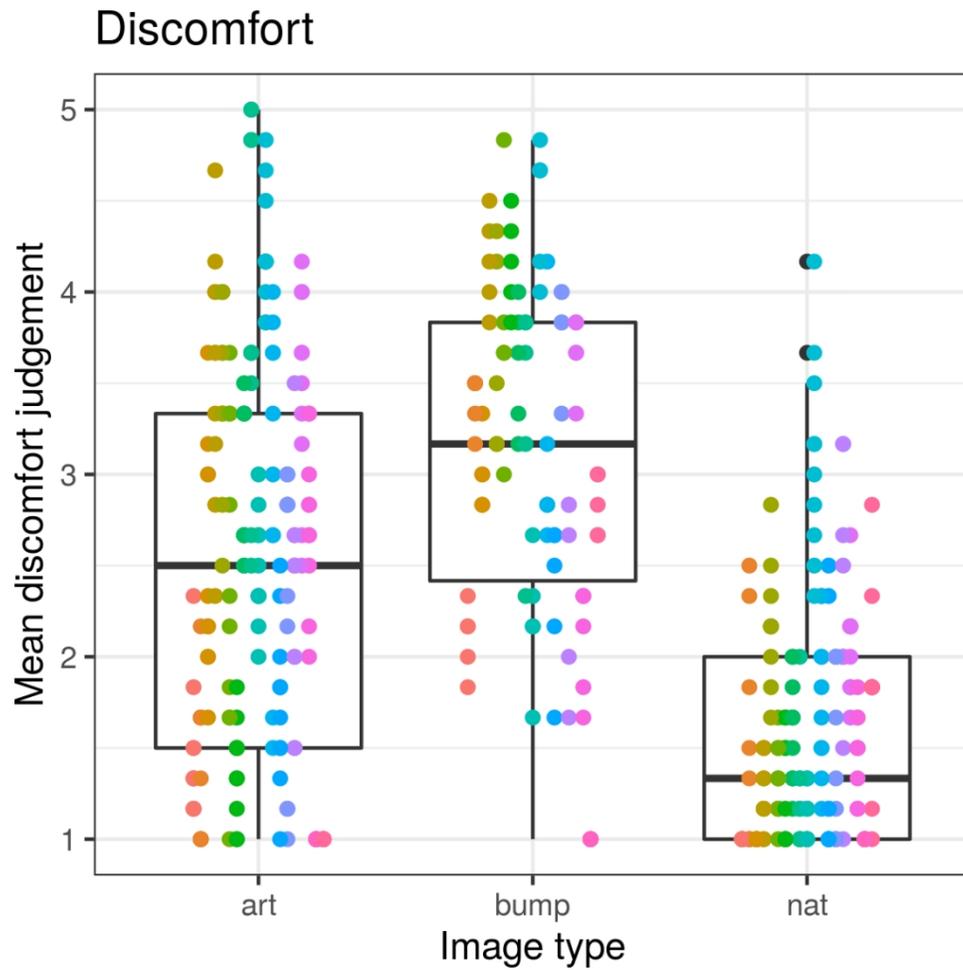


Figure 18 shows the average discomfort judgements (the mean from each individual) for the three image categories in Experiment 3. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

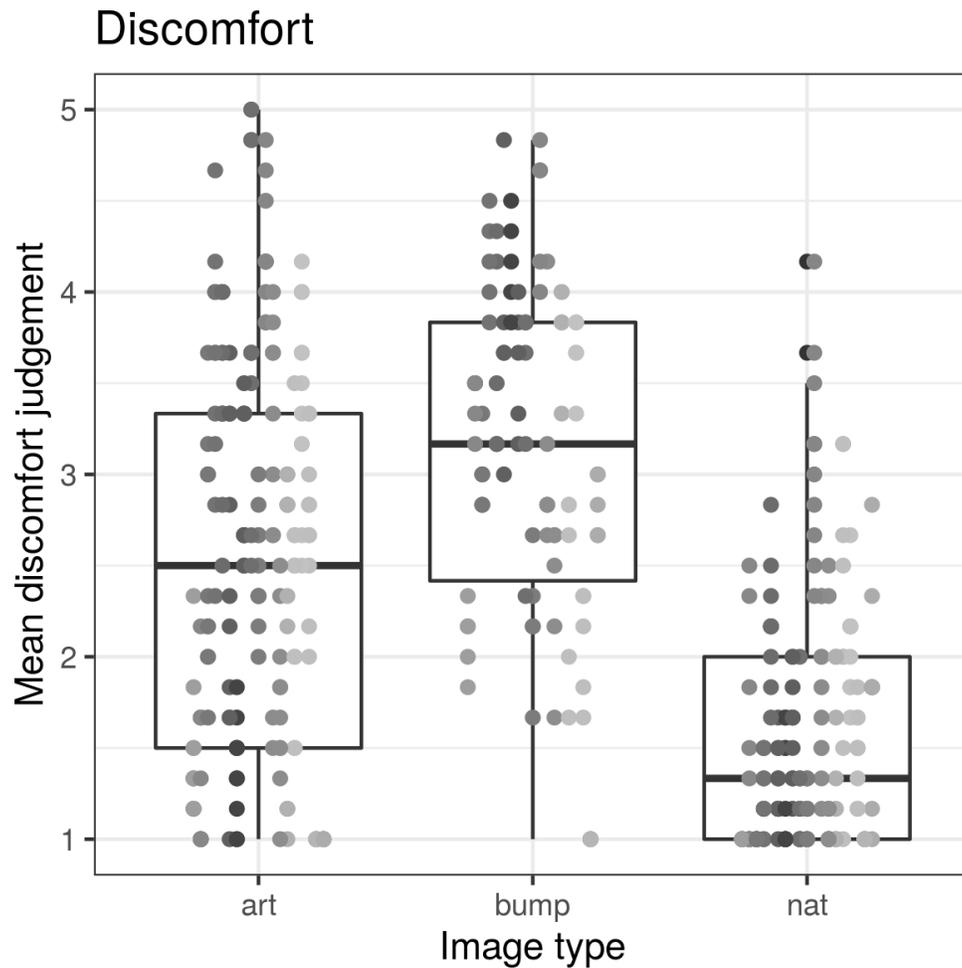


Figure 18 shows the average discomfort judgements (the mean from each individual) for the three image categories in Experiment 3. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

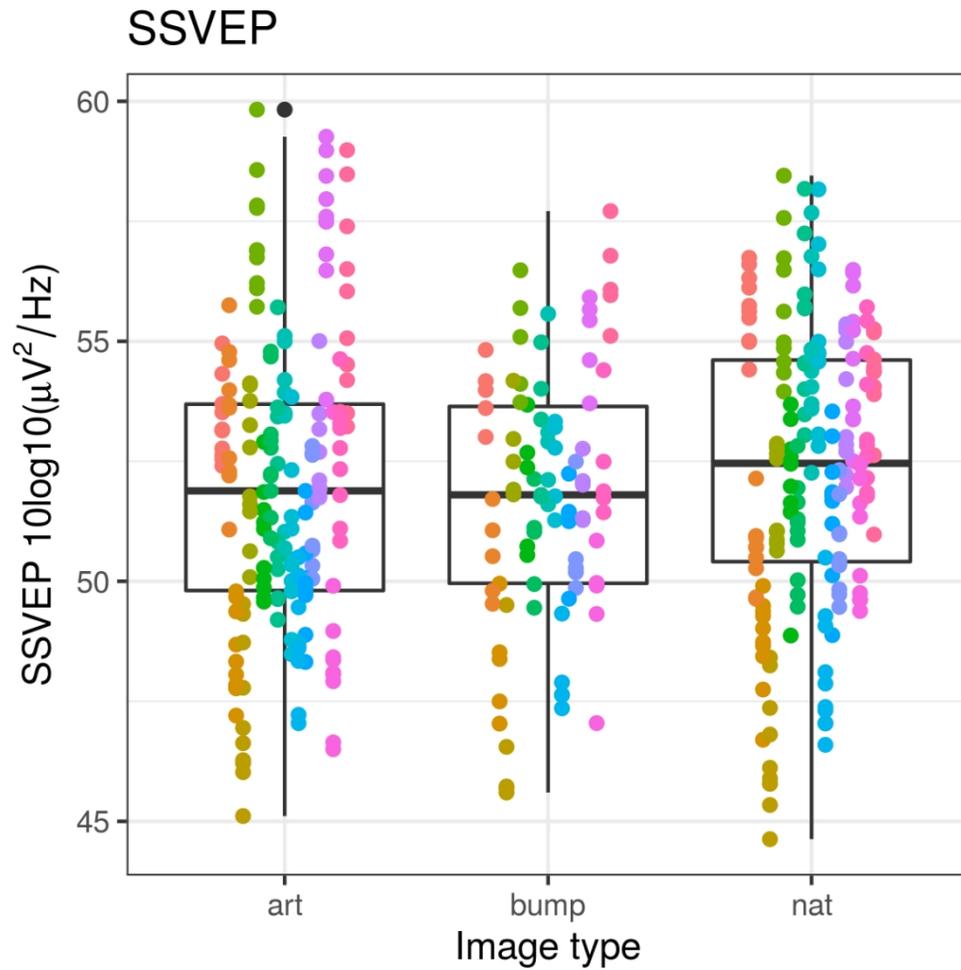


Figure 19 shows the SSVEP response to the three image categories in Experiment 3. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

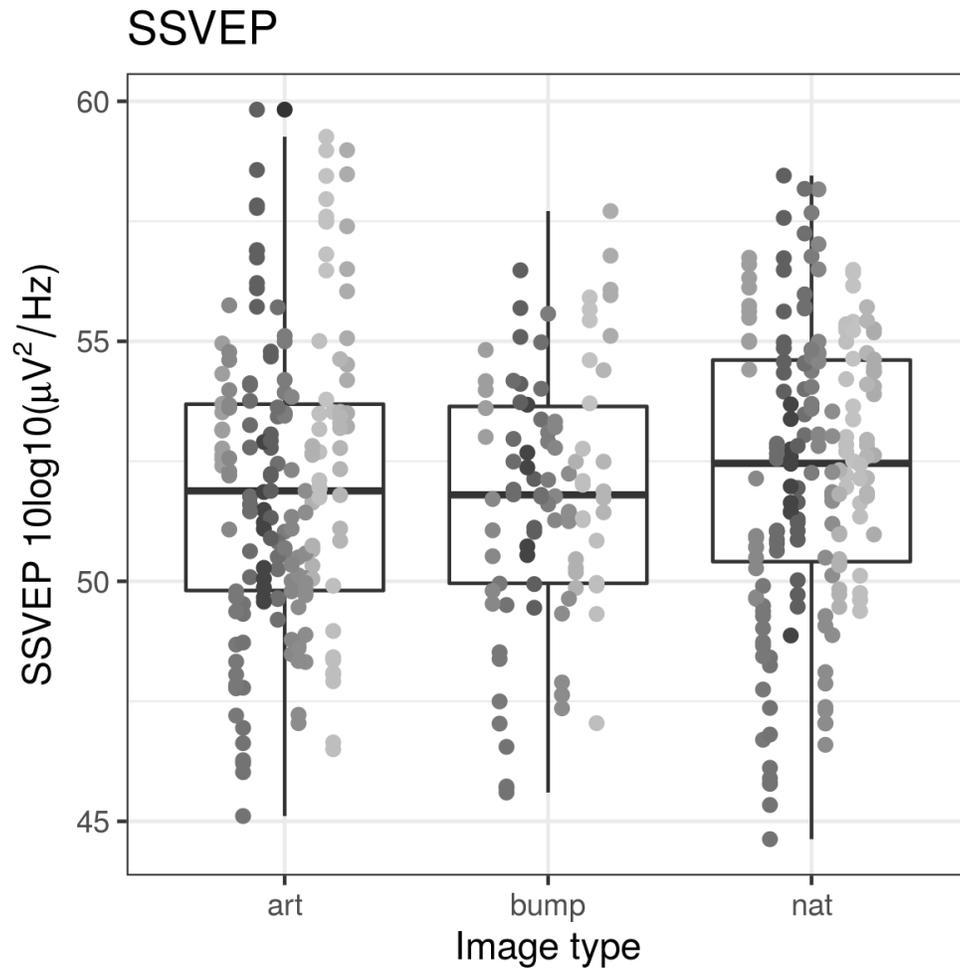


Figure 19 shows the SSVEP response to the three image categories in Experiment 3. Box boundaries indicate the 25th and 75th quartiles, whiskers indicate 1.5 times the inter-quartile range from the quartiles. Outliers are indicated as separate points (in black). Points superimposed on the plot represent individual observer responses.

101x101mm (300 x 300 DPI)

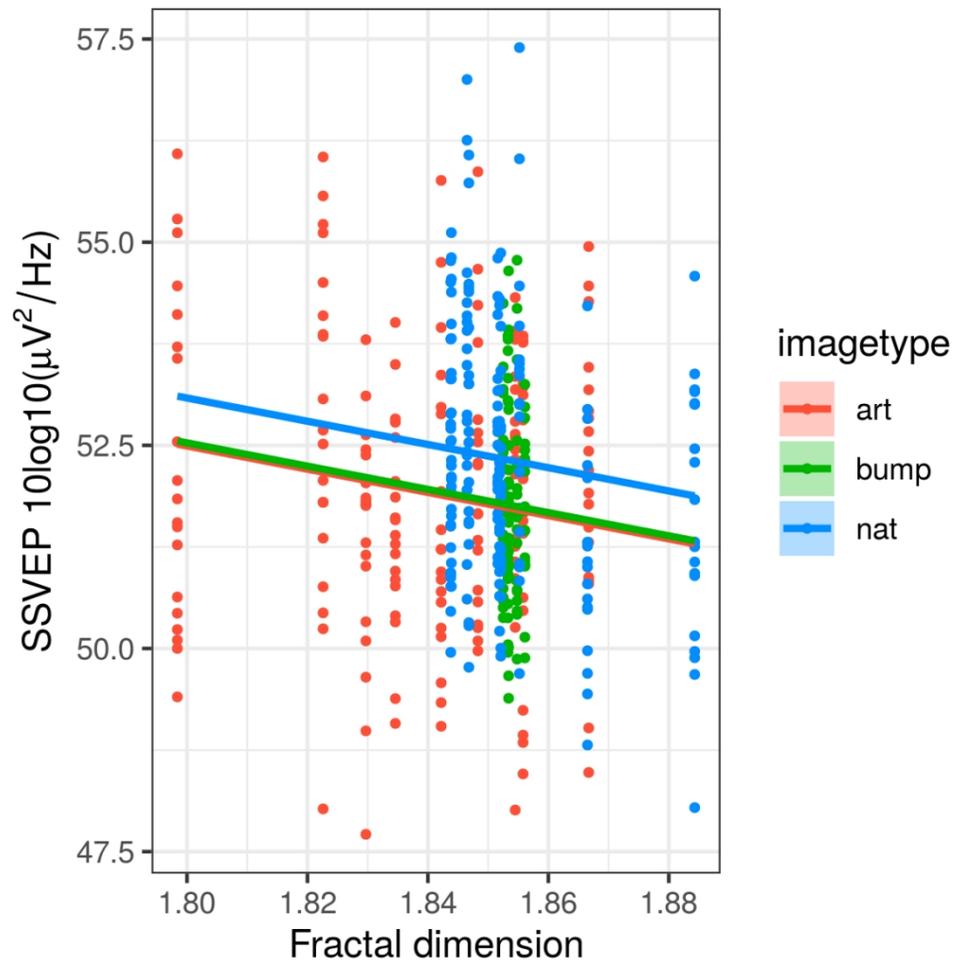


Figure 20 shows the model predicted SSVEP from fractal dimension. Image type was included as a fixed effect, and observer as a random effect.

101x101mm (300 x 300 DPI)

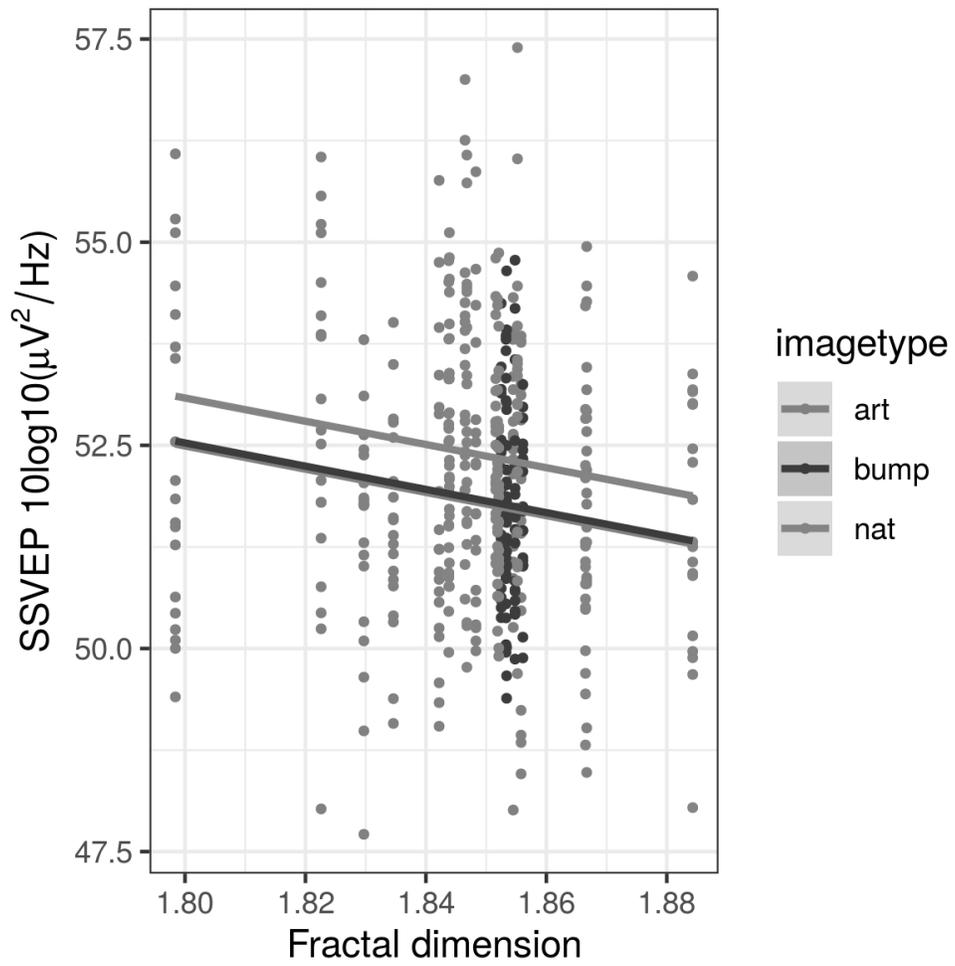


Figure 20 shows the model predicted SSVEP from fractal dimension. Image type was included as a fixed effect, and observer as a random effect.

101x101mm (300 x 300 DPI)

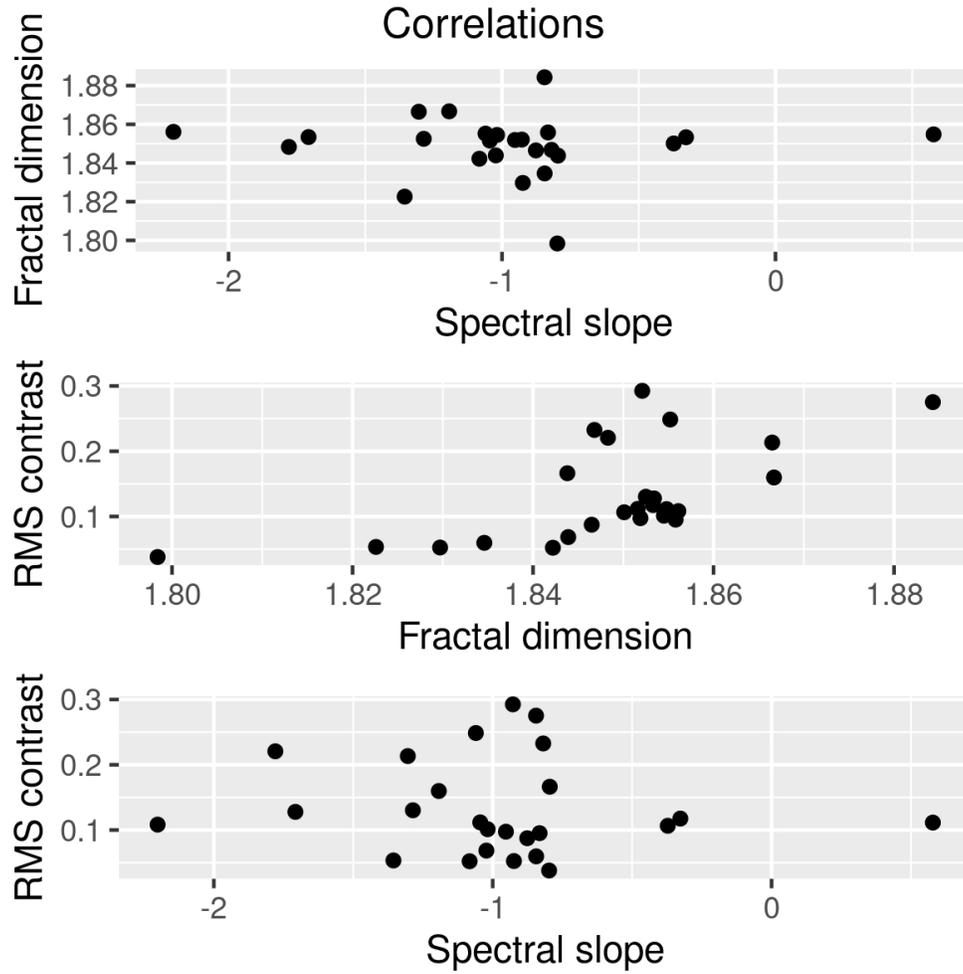


Figure 21 shows the relationship between (top) spectral slope and fractal dimension, (middle) fractal dimension and RMS contrast, and (bottom) spectral slope and RMS contrast for the images used in Experiment 3.

101x101mm (300 x 300 DPI)

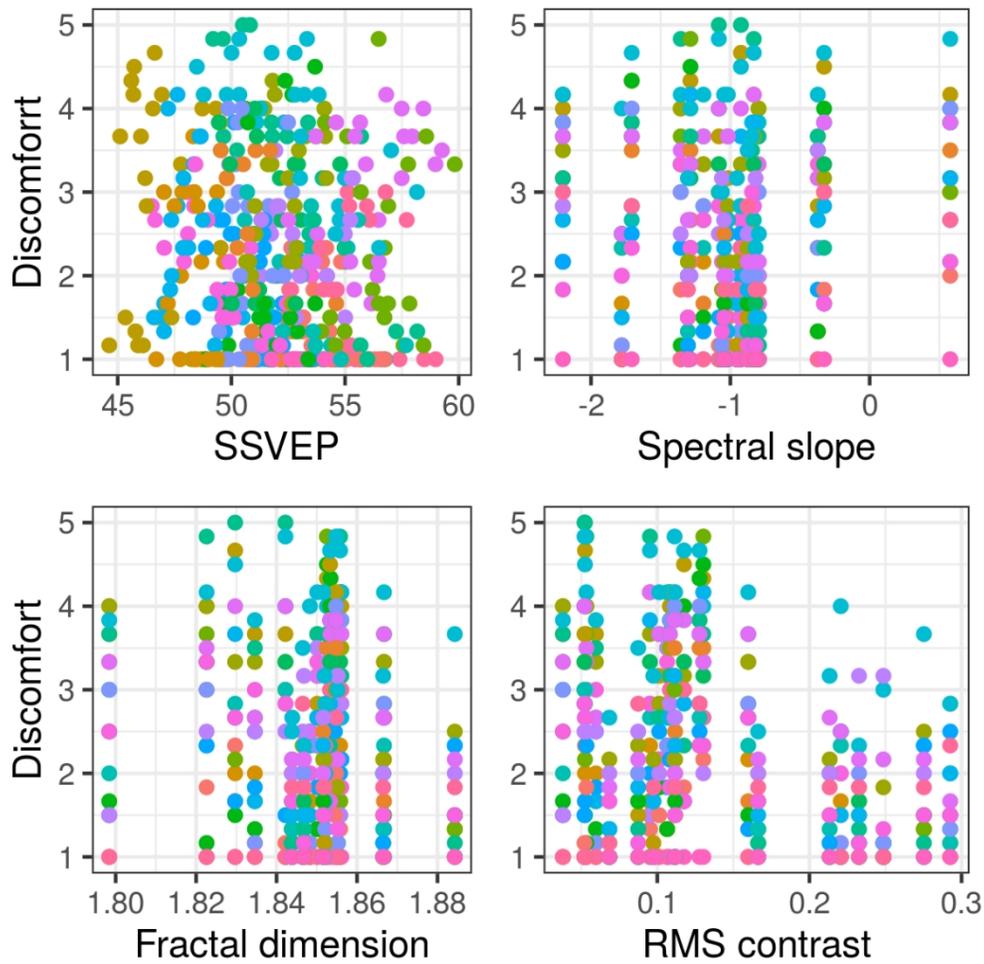


Figure 22 shows the bivariate correlations for each of the predictors included in the model against discomfort judgements. Top row shows SSVEP and spectral slope, bottom row shows fractal dimension and RMS contrast. Each individual is represented with a different shade.

101x101mm (300 x 300 DPI)

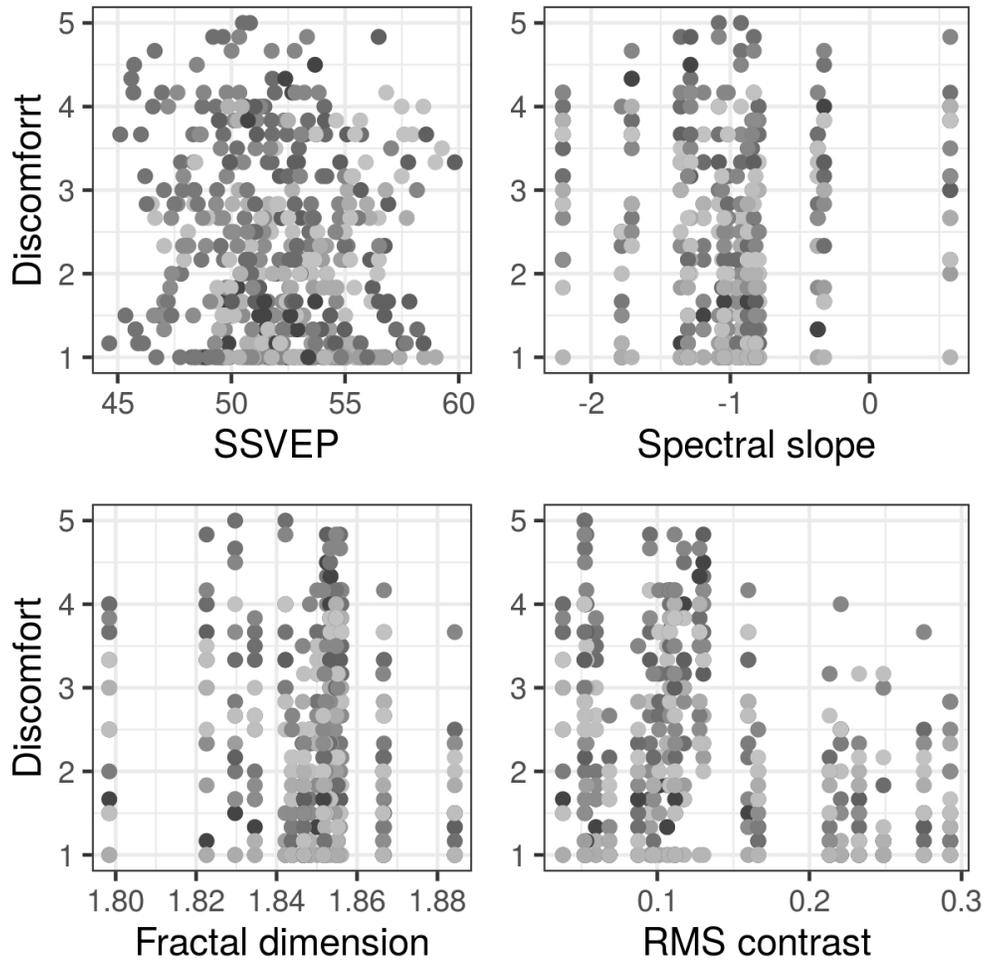


Figure 22 shows the bivariate correlations for each of the predictors included in the model against discomfort judgements. Top row shows SSVEP and spectral slope, bottom row shows fractal dimension and RMS contrast. Each individual is represented with a different shade.

101x101mm (300 x 300 DPI)