Vision Research 181 (2021) 47–60

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

# Nice and slow: Measuring sensitivity and visual preference toward naturalistic stimuli varying in their amplitude spectra in space and time

Zoey J. Isherwood [a,b,c,*], Colin W.G. Clifford [a], Mark M. Schira [b,d], Michelle M. Roberts [a,b], Branka Spehar [a]

[a] School of Psychology, UNSW Sydney, Sydney, NSW 2052, Australia
[b] School of Psychology, University of Wollongong, Wollongong, NSW 2522, Australia
[c] Department of Psychology, University of Nevada, Reno, NV 89557, USA
[d] Neuroscience Research Australia, Randwick, NSW 2031, Australia

ARTICLE INFO

*Keywords:*
Natural scene statistics
Spatiotemporal
Amplitude spectrum

ABSTRACT

The $1/f^{\alpha}$ amplitude spectrum is a statistical property of natural scenes characterising a specific distribution of spatial and temporal frequencies and their associated luminance intensities. This property has been studied extensively in the spatial domain whereby sensitivity and visual preference overlap and peak for slopes within the natural range ($\alpha \approx 1$), but remains relatively less studied in the temporal domain. Here, we used a 4AFC task to measure sensitivity and a 2AFC task to measure visual preference and across a wide range of spatial ($\alpha = 0.25$, 1.25, 2.25) and temporal ($\alpha = 0.25$ to 2.50, step size: 0.25) slope conditions. Stimuli with a shallow temporal slope modulate rapidly (e.g. 0.25), whereas stimuli with a steep slope modulate slowly (e.g. 2.25). Interestingly, sensitivity and visual preference did not closely overlap. While the sensitivity of the visual system is highest for our stimulus with an intermediate modulation rate (1.25), which is most abundant in nature, the stimulus with the slowest modulation rate (2.25) was most preferred. It seems sensible for the visual system to be sensitive to spatiotemporal spectra that most commonly exist in nature ($\alpha \approx 1$). However, it is possible that preference might be related to what these properties signal in the natural world. Consider the cases of waves slowly vs. rapidly crashing on a beach or fast vs. slow animals. In both instances the slowest option is often the safest and preferential, suggesting that the temporal $1/f^{\alpha}$ amplitude spectrum provides additional information that may indicate preferred environmental conditions.

## 1. Introduction

A number of studies have investigated the $1/f^{\alpha}$ amplitude spectrum in the *spatial* domain in an attempt to characterise how spatial regularities that typically exist across natural scenes influence the tuning of the visual system (Field, 1987; Knill et al., 1990; Tolhurst et al., 1992; Tadmor & Tolhurst, 1994; Webster & Miyahara, 1997; Párraga et al., 2000; Olman et al., 2004; Hansen & Hess, 2006; Spehar et al., 2015; Isherwood et al., 2017; Flitcroft et al., 2020). The $1/f^{\alpha}$ amplitude spectrum has also been characterised in multiple studies of visual preference, finding not only heightened sensitivity but also preference toward natural $1/f^{\alpha}$ spectra (Spehar et al., 2003, 2015, 2016; Juricevic et al., 2010; O'Hare & Hibbard, 2011; Spehar & Taylor, 2013; Penacchio & Wilkins, 2015; Viengkham et al., 2019; Nguyen & Spehar, 2021).

Following the discovery that static images of natural scenes follow a $1/f^{\alpha}$ distribution of luminance (Kretzmer, 1952; Burton & Moorhead, 1987; Field, 1987), theoretical investigations have been conducted to determine whether a $1/f^{\alpha}$ distribution of frequencies also exists as a function of *time* (Eckert & Buchsbaum, 1993; Van Hateren, 1993). This theoretical work was soon followed up with strong empirical evidence by Dong and Atick (1995), and later by Billock et al. (2001b), that movies of natural scenes indeed follow a $1/f^{\alpha}$ distribution of luminance in both space *and* time ($\alpha \approx 1$). Yet, since the discovery of this *temporal* property of natural scenes, only a few studies have attempted to characterise the sensitivity of the visual system to variations in *spatiotemporal* $1/f^{\alpha}$ spectra (Billock et al., 2001a; Baker & Graf, 2009; Cass et al., 2009).

The $1/f^{\alpha}$ amplitude spectrum in the *temporal* domain can be manipulated in the same way as the *spatial* domain—through changes in $\alpha$—which will henceforth be referred to as changing either the *spatial slope* (SS) or the *temporal slope* (TS) of a stimulus. Similar to how

* Corresponding author at: Department of Psychology, University of Nevada, Reno, Reno, NV 89557, USA.
*E-mail address:* zisherwood@unr.edu (Z.J. Isherwood).

https://doi.org/10.1016/j.visres.2021.01.001
Received 15 July 2020; Received in revised form 6 January 2021; Accepted 6 January 2021
Available online 9 February 2021

changing the SS changes the degree in which certain SF bands are perceived in an image (e.g. a *shallow* SS results in an image with enhanced edges, while a *steep* SS results in a "blurred" image—see Fig. 1A), manipulating the TS of a dynamic stimulus changes the degree in which certain TF bands are perceived. For instance, if a stimulus has a *shallow* TS it appears as flickering or the rate of luminance modulation is *rapid* due to the large amount of energy contained in the high TF domain (see Fig. 1B). A *steep* TS, on the other hand, is perceived as having a rate of luminance modulation that changes *slowly* due to the large amount of energy contained in the low TF domain. For examples of stimuli across different SS and TS conditions, see Movie 1 (corresponding to Fig. 1B, https://osf.io/3q8gm/) and Movie 2 (corresponding to Fig. 2, https://osf.io/qbr3z/).

Billock et al. (2001a) used this manipulation to investigate visual tuning toward the $1/f^{\alpha}$ amplitude spectrum in both space and time using synthetic noise movies similar to those depicted in Movie 2. To solely assess the spatiotemporal $1/f^{\alpha}$ amplitude spectrum, random noise stimuli are used as they can be highly controlled for low-level properties and are devoid of other higher-order natural image properties such as those present in Fig. 1 and Movie 1 (e.g. in phase intact images edge positions are dependent on objects in a natural scene), which may influence perceptual judgments (Yoonessi and Kingdom, 2008). Billock et al. (2001a) found that just noticeable difference (JND) thresholds were lowest for SS and TS combinations that were most *natural* ($\alpha \approx 1$), suggesting that the visual system is sensitive to natural $1/f^{\alpha}$ spectra in space and time. In addition, JND thresholds were lowest when stimuli had $1/f^{\alpha}$ spectra that were *concordant* in space and time (e.g. SS 1.25 TS 1.25)—irrespective of whether or not the SS or TS of a stimulus was *natural*.

These findings may reflect that the spatial and temporal components of the $1/f^{\alpha}$ amplitude spectrum are generally *inseparable* across natural scenes (Dong & Atick, 2009). Indeed, this space–time inseparability can be observed in Movie 1 (corresponding to Fig. 1B, https://osf.io/3q8gm/). Changes to the temporal $1/f^{\alpha}$ spectrum of a natural movie (such as the one depicted in Movie 1) result in changes to the movie's spatial $1/f^{\alpha}$ spectrum—pulling it in the direction of the manipulation (e.g. *steepening* the temporal slope of a natural movie will also make its spatial slope *steeper*). This is not the case, however, for synthetic noise stimuli which can be generated to be space–time separable—see Movie 2 for examples (corresponding to Fig. 2, https://osf.io/qbr3z/).

Few studies have characterised the tuning of the visual system toward the *spatiotemporal* $1/f^{\alpha}$ amplitude spectrum (Baker & Graf, 2009; Cass et al., 2009). As such, visual sensitivity to this property is currently unknown beyond the range Billock et al. (2001a) used, where SS varied from 0.4 to 2.2 and TS varied from 0.2 to 1.4. It also remains unclear

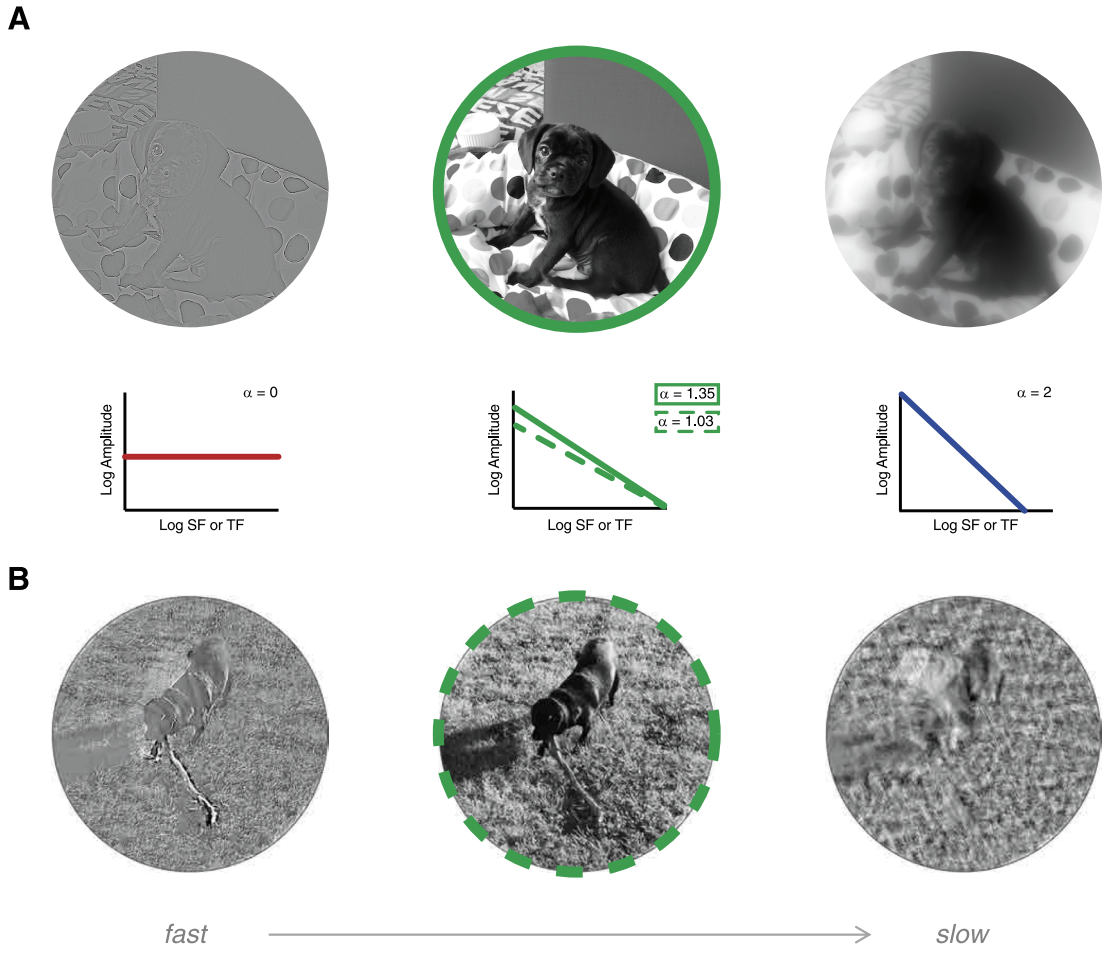


**Fig. 1.** Examples of spatial (A) and temporal (B) manipulations of the $1/f^{\alpha}$ amplitude spectrum in a real-world image and movie. For the corresponding movie for (B), see Movie 1 (https://osf.io/3q8gm/). A) The picture depicted in the middle panel has a SS ($\alpha = 1.35$, solid line in graph) within the natural range. A shallower SS (e.g. $\alpha = 0$) enhances edges due to an increase of energy across high SF bands, and a decrease across low SF bands (leftmost panel). A steeper SS (e.g. $\alpha = 2$) causes the image to look blurry due to an increase in energy across low SF bands and a decrease across high SF bands (rightmost panel). B) The movie depicted in the middle panel has a TS ($\alpha = 1.03$, dashed line in graph) within the natural range. Making the TS shallower (e.g. $\alpha = 0$) causes only the fast components of the movie to be perceived due an increase in energy across high TF bands and a decrease across low TF bands—any rapid movement (such as the dog chewing the stick) corresponds to heightened luminance (leftmost panel). Making the TS steeper (e.g. $\alpha = 2$) causes only the slow components of the movie to be perceived due to an increase in energy across low TF bands and a decrease across high TF bands (rightmost panel).

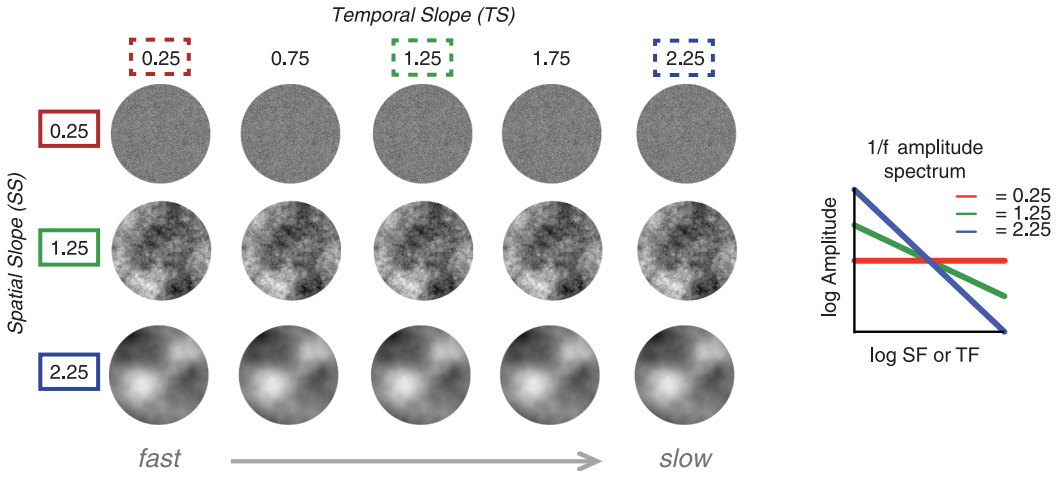


**Fig. 2.** Examples of stimuli used in the present study. Synthetic noise movies were generated across 3 spatial slope (rows) and 5 temporal slope (columns) conditions at 30% RMS contrast, which resulted in 15 unique combinations. Shallow temporal slopes appear to have rapid luminance modulation (e.g. TS 0.25), whereas steep temporal slopes appear to have slow luminance modulation (e.g. TS 2.25). For the corresponding movie (Movie 2), see the following link: https://osf.io/qbr3z/.

whether sensitivity toward spatiotemporal variations in the amplitude spectrum correlate with visual preference. In the *spatial* domain, there appears to be a tight link between visual preference and sensitivity and it has been suggested that preference toward natural statistical properties may be driven by the tuning properties of the visual system (Reber et al., 2004). Visual preference toward variations in *spatial* $1/f^{\alpha}$ spectra is well established, where peak preference is found for stimuli with a natural SS, with lowered preference found for shallower and steeper SS conditions (Juricevic et al., 2010; Spehar et al., 2015, 2016). This inverse U-shaped response profile is also observed when measuring sensitivity across SS conditions (Knill et al., 1990; Tadmor & Tolhurst, 1994; Spehar et al., 2015) and is observed for BOLD responses in early visual cortex (Isherwood et al., 2017; Olman et al., 2004). It is currently unknown, however, the full extent to which these associations also exist in the *temporal* domain (Juricevic et al., 2010; Spehar & Taylor, 2013; Spehar et al., 2015; Yoshimoto et al., 2017, 2019, 2020).

To address these gaps, we conducted a series of psychophysics experiments measuring sensitivity and visual preference to a range of spatial and temporal $1/f^{\alpha}$ spectra. In *Experiment 1*, participants were presented random noise movies varying in their spatial ($\alpha = 0.25, 1.25, 2.25$) and temporal ($\alpha = 0.25, 0.75, 1.25, 1.75, 2.25$) $1/f^{\alpha}$ amplitude spectra across two different task conditions: 1) sensitivity—a four alternative choice (4AFC) task where participants selected the "odd stimulus out," providing a measure of discrimination sensitivity and 2) visual preference—a two alternative forced choice (2AFC) task where participants indicated their preference between two stimuli. In *Experiment 2* we tested whether the presentation order of our stimuli would affect their measured visual preference. For this we compared our visual preference measures from *Experiment 1*, which had an intermixed presentation order, with visual preference measured using a fixed presentation order (SS fixed across blocks). Lastly, in *Experiment 3* we measured sensitivity and visual preference to a different range of TS values ($\alpha = 0.50, 1.00, 1.50, 2.00, 2.50$) across the same SS conditions used in *Experiment 1* and *2* ($\alpha = 0.25, 1.25, 2.25$). This experiment tested the generalisability of our findings in *Experiment 1* to a different sample, as well as extending our measures across a wider range of TS conditions.

We predicted that the most *natural* stimulus in our set, both spatially and temporally, would be the easiest to discriminate and the most preferred. We also predicted that the same pattern of results would be observed for visual preference and sensitivity measures across all other SS and TS conditions. If our predictions are supported, this would provide evidence that the visual system has evolved to become tuned to *both* the spatial and temporal statistical properties of natural scenes. This would also support the notion that visual preference toward statistical properties that exist in nature is driven by the ease in which these properties are processed by the visual system (Reber et al., 2004).

## 2. Experiment 1: Measuring sensitivity and visual preference toward spatiotemporal 1/fα amplitude spectra

### 2.1. Methods

#### 2.1.1. Design

A 2 (task) × 3 (spatial slope, SS) × 5 (temporal slope, TS) repeated measures design was used. The task conditions were: 1) sensitivity and 2) preference. The SS conditions used were: 0.25, 1.25, 2.25. The TS conditions used were: 0.25, 0.75, 1.25, 1.75, 2.25.

#### 2.1.2. Apparatus

Stimuli were presented and generated using MATLAB (2012a, version 7.14) software (MathWorks) and Psychophysics Toolbox functions (Brainard, 1997; Pelli, 1997), and also driven by a Bits# stimulus processor (Cambridge Research Systems, Rochester, Kent, UK), which provides 14-bit grayscale resolution. A gamma-corrected 18-inch ViewSonic Graphics Series G90f CRT monitor (resolution, 1280 × 1024) was used to present stimuli at 85 Hz. The experiment was conducted in a dark cubicle, and participants used a chin rest to maintain a viewing distance of 35 cm. Responses were collected using a regular computer keyboard.

#### 2.1.3. Participants

33 participants with normal or corrected-to-normal vision were recruited using the University of New South Wales (UNSW) Psychology SONA system. SONA participants consisted of students enrolled in a first-year psychology elective (PSYC1001) at UNSW who were required to participate in experiments for course credit. Two participants were excluded from analysis due to non-compliance with task instructions. As such, data from 31 participants was analysed. The number of participants recruited for the present study was motivated by our previous work (Spehar et al., 2003, 2015, 2016; Spehar & Taylor, 2013). Ethics approval was provided by the UNSW Human Research Advisory Panel (Reference Number: HREAP-C 2349).

#### 2.1.4. Stimuli

Synthetic noise movies were generated using a custom code in MATLAB that allows precise control of the distribution of pixel intensities (1/f slope) and root mean squared (RMS) contrast. Each stimulus was created starting from a normally distributed ($\mu = 0$, $\sigma = 1$)

matrix of random noise ($A_{noise}$) generated at a desired size (128, 128, 128) across spatial dimensions ($x$ & $y$) and frames ($z$). The noise matrix was then 3D fast Fourier transformed (FFT), and then the DC component was subtracted to remove DC bias. A transformation matrix ($A_{ampspec}$) proportional to $1/f^{\alpha}$ based on an input α in space ($\alpha_{spatial}$) and an input α in time ($\alpha_{temporal}$) was subsequently generated. This matrix was generated to have the same dimensions as $A_{noise}$. Values in this matrix were generated by first calculating the polar distance of each point in the matrix across spatial ($d_{spatial} = \sqrt{(x-x_0)^2 + (y-y_0)^2}$) and temporal ($d_{temporal} = \sqrt{(z-z_0)^2}$) dimensions in Fourier space. Polar distance values were then raised to an exponent that corresponded to the desired input α in space and time, and subsequently multiplied together across dimensions ($A_{ampspec} = d_{spatial}^{(1/\alpha_{spatial})} * d_{temporal}^{(1/\alpha_{temporal})}$). The transformation matrix was then multiplied in Fourier space with the noise matrix ($A_{ampspec} * A_{noise}$) and inverse Fourier transformed—yielding a stimulus with a specified $1/f^{\alpha}$ spatiotemporal amplitude spectrum. The MATLAB script used to generate the stimuli was make_fractal_3D.m, which can be viewed at the following link: https://osf.io/w5tvn/.

For the visual preference task, 3 different seeds of random noise were used to generate dynamic stimuli at 30% RMS contrast across three different SS values (α = 0.25, 1.25, 2.25) and five different TS values (α = 0.25, 0.75, 1.25, 1.75, 2.25) resulting in 15 unique combinations (Fig. 2). The size of each stimulus was 128 × 128 × 128 pixels, subtending 5.48° visual angle.

For the discrimination sensitivity task, the Psi adaptive staircase was used (Kontsevich & Tyler, 1999) with 78 possible steps (39 for incremental trials and 39 for decremental trials), resulting in 1,170 possible TS increments and decrements which were pre-generated for the experiment to meet computational demand. The same stimulus parameters as in the visual preference task were used (contrast, size, α values etc.) across all possible combinations.

When generating stimuli with a specific input α value there are some slight discrepancies between the input and output α value caused by rounding to the nearest 14-bit integer (see Tables 1 and 2). All terminology referring to SS and TS will refer to the input α value unless otherwise stated.

The spatial α of each stimulus was measured by plotting the $1/f^{\alpha}$ amplitude spectrum ($A(sf)$) on a *linear* axis using MATLAB. This was done by first conducting a 2D FFT of each frame of the stimulus, shifting the zero-frequency component to the centre of the spectrum, finding the magnitude of the transform, and lastly averaging across $x$ & $y$ dimensions. After plotting the spatial $1/f^{\alpha}$ amplitude spectrum of the stimulus, the data was fit on a linear axis as a function of SF for each frame (1 2 8) separately after 14-bit conversion. A non-linear fit function in MATLAB with a multiplicative inverse model ($\frac{1}{x^{\alpha}}$) was used to fit the data with starting coefficients of −1 as a minimum, and the maximum value of $A(sf)$ as a maximum, yielding an estimated spatial α for each frame. The measured spatial α reported in Table 1 was averaged across frames and three different spatial patterns (seeds); the standard deviation is reported in parentheses. The MATLAB script used to measure the spatial α of our stimuli was calc_spatialslope.m, which can be viewed at the following link: https://osf.io/98gkt/.

**Table 1**
Measured SS across each SS × TS combination.

| | | Input TS | | | | |
|---|---|---|---|---|---|---|
| | | 0.25 | 0.75 | 1.25 | 1.75 | 2.25 |
| Input SS | 0.25 | 0.22 (0.03) | 0.23 (0.03) | 0.23 (0.03) | 0.22 (0.02) | 0.22 (0.02) |
| | 1.25 | 1.20 (0.09) | 1.20 (0.09) | 1.19 (0.10) | 1.17 (0.10) | 1.17 (0.10) |
| | 2.25 | 2.12 (0.29) | 2.08 (0.26) | 2.10 (0.33) | 2.10 (0.35) | 2.06 (0.33) |

*Notes:* Reported SS values were averaged across 3 seed conditions and 128 frames (384 frames total). Values in parentheses indicate the standard deviation. Fit was conducted in linear space.

**Table 2**
Measured TS across each SS × TS combination.

| | | Input TS | | | | |
|---|---|---|---|---|---|---|
| | | 0.25 | 0.75 | 1.25 | 1.75 | 2.25 |
| Input SS | 0.25 | 0.25 (0.00) | 0.75 (0.00) | 1.25 (0.00) | 1.76 (0.00) | 2.24 (0.00) |
| | 1.25 | 0.25 (0.00) | 0.75 (0.00) | 1.26 (0.01) | 1.76 (0.02) | 2.20 (0.11) |
| | 2.25 | 0.24 (0.02) | 0.73 (0.03) | 1.26 (0.04) | 1.73 (0.02) | 2.38 (0.19) |

*Notes:* Reported TS values were averaged across 3 seed conditions. Values in parentheses indicate the standard deviation. Fit was conducted in linear space.

The measured temporal α reported in Table 2 was averaged across three different spatial patterns (seeds), and the standard deviation across seeds is reported in parentheses. The temporal α of each stimulus was measured by plotting the temporal $1/f^{\alpha}$ amplitude spectrum ($A(f)$) on a *linear* axis using MATLAB functions. This was done by first conducting a 3D FFT in MATLAB, collapsing across $x$ and $y$ dimensions, and removing singleton dimensions, resulting in an array of estimated amplitude across the $z$ dimension—i.e. luminance intensity over time. These data were subsequently plotted in linear space as a function of TF after 14-bit conversion (Fig. 3, middle panel). A non-linear fit function in MATLAB with a multiplicative inverse model ($\frac{1}{x^{\alpha}}$) was used to fit the data, with starting coefficients of −1 as a minimum, and the maximum value of $A(tf)$ as a maximum, yielding an estimated temporal α. The MATLAB script used to measure the temporal α of our stimuli was calc_temporalslope.m, which can be viewed here: https://osf.io/bz5hd/.

We used *linear* space rather than log–log space (the traditional way of fitting the $1/f^{\alpha}$ amplitude spectrum) as it is more robust to changes caused by 14-bit conversion (e.g. the flattening of the amplitude spectrum across high TFs; see blue circles in the bottom panel of Fig. 3). In linear space these changes are minor, however in log–log space they are emphasised (particularly for *temporal* $1/f^{\alpha}$ spectra)—which consequently affects the fit.

### 2.1.5. Procedure

*2.1.5.1. Discrimination Threshold Measurements.* To determine the discrimination threshold contrast (JND) necessary to detect increases and decreases in the SS and TS of the reference image, the Bayesian adaptive Psi procedure was used (Kontsevich & Tyler, 1999; Spehar et al., 2015). We employed a four alternative forced choice (4AFC) "odd-one-out" task, in which the observer was asked to indicate which stimulus was different among the four stimuli shown on any given trial. This design has been demonstrated by Jakel and Wichmann (2006) as well as Vancleef et al. (2018) to provide a more robust measurement of thresholds in fewer trials compared to 2AFC or 2IFC tasks in naïve observers.

Each trial began with a fixation point (4 × 4 pixels, 0.17° visual angle) at the centre of the screen for 1 s, followed by a trial display in which four stimuli were shown for a period of 2.133 s (not looped as for the visual preference task). Each stimulus was 192 pixels (8.21° visual angle) apart, and presented in a circular aperture with a blurred (raised cosine) edge.

SS and TS conditions were not intermixed. Each run of the experiment consisted of 1 SS and a base TS which was either increased or decreased in a trial depending on the participants' previous responses. In a single trial, all four stimuli had the same SS and 3 of the stimuli had the same TS. The remaining odd-one-out stimulus had a different TS, which appeared randomly with equal probability in each of the four quadrants. Each run consisted of 60 trials, 30 where the odd-one-out stimulus TS

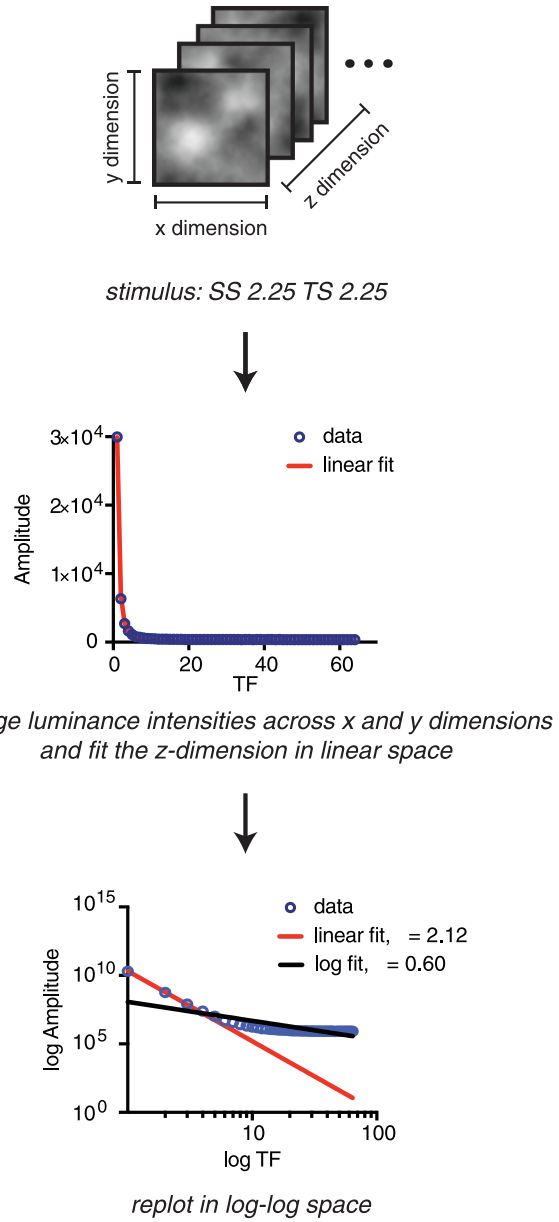


**Fig. 3.** A depiction of how the temporal $1/f^{\alpha}$ spectrum of each stimulus was measured. The example depicted here is for a stimulus with a SS and TS of 2.25 (top panel). First a 3D Fourier Transform was conducted on the stimulus, then luminance intensity (amplitude) was collapsed spatially across the x and y dimensions of the stimulus in Fourier space. The temporal (z) dimension was then plotted on a linear axis where amplitude was fitted as a function of temporal frequency (TF) (middle panel). The temporal $1/f^{\alpha}$ amplitude spectrum of our stimuli were fitted in linear space in order to account for changes caused by 14-bit conversion, which are emphasised when plotted on a log–log axis (e.g. the flattening of the amplitude spectrum at high TFs) (bottom panel). The TS fit on a linear axis is much closer to the input TS (2.12 v 2.25) compared to the log–log fit (0.60 v 2.25).

increased, and 30 where the odd-one-out stimulus TS decreased. Incremental (Up) and decremental (Down) trials were intermixed. The 15 conditions of SS and TS combinations resulted in a total of 900 trials. Including optional breaks at the end of each run, the total time to complete this task was around 45 min. See Fig. S11 in the Supplementary Materials for a schematic of the procedure.

Due to computational limitations, we were unable to present 4 different seed patterns (i.e. 4 stimuli with different noise patterns) each trial with accurate timing. Instead, in each trial the three distractor stimuli were the same stimulus—but rotated relative to one another to avoid repeating the same pattern with the same orientation. Since the stimuli were dynamic random noise movies, these 3 distractor stimuli were essentially indistinguishable (see Movie S-1, https://osf.io/gaqdb/), enabling accurate stimulus presentation each trial. The task design facilitated focus on the overall appearance of the stimulus rather than adopting an image matching strategy. The duration of the response interval was unlimited. Participants were given auditory feedback, with either one or two beeps indicating the response was correct or incorrect, respectively.

*2.1.5.2. Visual Preference Measurements.* Visual preference was measured using a two-alternative forced-choice (2AFC) paired comparison procedure (Cohn, 1894; Spehar et al., 2015). Each trial began with a fixation point (12 × 12 pixels, 0.51° visual angle) at the centre of the screen for 500 ms. Subsequently, two images were presented side-by-side with observers indicating (via key press) which of the two stimuli they visually prefer. To avoid imposing a criterion for the subject to use during this task, no further instruction was given other than to "pick the noise stimulus that you prefer". Stimuli were presented 200 pixels (8.55° visual angle) apart in a circular aperture with a blurred (raised cosine) edge. The duration of the response interval was unlimited. Due to the periodic nature of the Fourier transform, there was no temporal discontinuity while the stimuli were looped until the response was recorded (Baker & Graf, 2009).

In this procedure, stimuli of each SS (3) and TS (5) combination (i.e. 15 different stimuli) are paired with all of the other SS and TS combinations in the series, resulting in a total of 210 pairs. This basic sequence ensures that every experimental image is paired with each other and that each experimental stimulus is presented 28 times across all experimental pairings with equal frequency on the left and the right side. All pairs were presented in random order. A complete sequence of 210 trials was only presented once with no repetitions across 3 blocks (70 trials/block). Participants were allocated breaks at the end of each block. This task took around 15 min for participants to complete. See Fig. S-2 in the Supplementary Materials for a schematic of the procedure.

### 2.2. Results

Initial analyses were conducted on responses from each task individually (sensitivity and visual preference). Subsequent analyses involved comparing responses across both tasks. For this, data were transformed into z-scores to allow a direct comparison between them. All statistics reported in this section were Greenhouse-Geisser corrected. All the raw data (sensitivity and visual preference) have been anonymised and are available on the following OSF project webpage: https://osf.io/6sgpy/.

#### 2.2.1. Sensitivity

A 2-way repeated measures ANOVA was conducted on the sensitivity (1/threshold) values obtained from the 4AFC discrimination task. Here, we report statistical analyses that were conducted on sensitivity values that were averaged across both *up* and *down* staircases of the task (i.e. increments and decrements), which is common practice when calculating psychophysical thresholds (Kingdom & Prins, 2016). A table summarising the main effects and interactions can be found in the Supplementary Materials (Table S-1).

#### 2.2.2. Main effects and interactions

The main effects of *SS* ($F_{1.629,48.868} = 17.953$, $p < 0.001$, $\eta_p^2 = 0.374$) and *TS* ($F_{1.571,47.153} = 64.843$, $p < 0.001$, $\eta_p^2 = 0.684$) were significant. The direction of these effects is depicted in Fig. 4 below, which show an inverted U-shaped pattern of results where sensitivity is highest for the most *natural* SS (Fig. 4A) and the most *natural* TS (Fig. 4B) in the stimulus set—1.25. There was a significant interaction between *SS* × *TS*

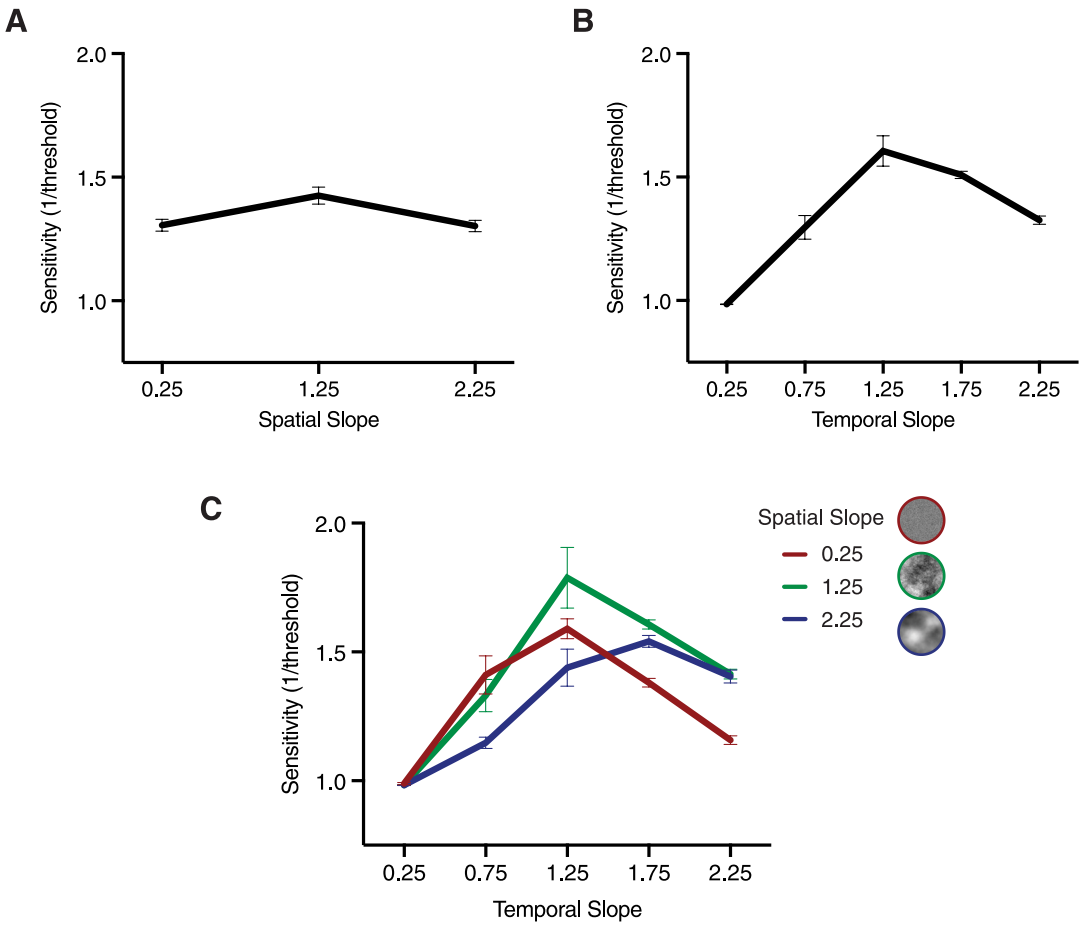


**Fig. 4.** Main effects and interactions from the sensitivity analysis. A) Main effect of SS—sensitivity plotted as a function of SS. The values plotted were averaged across all TS conditions. An inverted U-shaped pattern peaking for SS 1.25 is observed. B) Main effect of TS—sensitivity plotted as a function of TS. The values plotted were averaged across all SS conditions. An inverted U-shaped pattern peaking for TS 1.25 is observed. C) Interaction between SS and TS—sensitivity plotted as a function of TS for each SS condition separately (red, green, blue lines correspond to SS 0.25, SS 1.25, SS 2.25 respectively). Error bars depict SEM between subjects. See Fig. S-3 in the Supplementary Materials to view the same plots with individual datapoints. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

($F_{2.365,\ 70.958} = 11.065$, $p < 0.001$, $\eta_p^2 = 0.269$), and the direction of this interaction is plotted in Fig. 4C. The effect sizes reported here (as indicated by $\eta_p^2$) are considered large as per the effect size conventions of Cohen (1988).

Overall, the most *natural* stimulus in the set elicits the highest sensitivity (SS 1.25 TS 1.25). There is also a rightward shift in sensitivity from *shallow* to *steep* SS conditions. For SS 0.25, sensitivity is heightened between TS 0.75 and TS 1.25; for SS 1.25, sensitivity is heightened between TS 1.25; and 1.75, and for SS 2.25, sensitivity is heightened between TS 1.25 and TS 2.25. With the exception of TS 0.25 it appears that *the closer the spatial and temporal slope, the higher the sensitivity* (e.g. SS 1.25 TS 1.25).

### 2.2.3. Visual Preference

A 2-way repeated measures ANOVA was conducted on the *proportion chosen* values obtained from the 2AFC visual preference task. Before statistical analysis, the raw data were arcsine transformed. This transform was used as a variance-stabilisation measure as the raw data had a non-normal (binomial) distribution (McDonald, 2014). A table summarising the main effects and interactions can be found in the Supplementary Materials (Table S-2).

### 2.2.4. Main effects and interactions

The main effect SS was not significant ($F_{1.389,41.675} = 2.509$, $p = 0.110$, $\eta_p^2 = 0.077$), however TS was significant ($F_{1.177,35.305} = 26.735$, $p < 0.001$, $\eta_p^2 = 0.471$). The direction of these effects is depicted in Fig. 5A and B, showing that visual preference is highest for steep SS and steep TS conditions. These findings indicate that on average the most preferred stimulus in the set was SS 2.25, TS 2.25. There was also a significant interaction between *SS* × *TS* ($F_{4.585,137.564} = 15.827$, $p < 0.001$, $\eta_p^2 = 0.345$), which is depicted in Fig. 5C. All SS conditions show increasing preference as the TS steepens. Generally, across TS conditions SS 2.25 was most preferred (blue line in Fig. 5C), followed by SS 1.25 then SS 0.25. The effect size for SS is considered small to medium, and the effect sizes for TS and for the SS × TS interaction are considered large as per Cohen's (1988) effect size conventions.

### 2.2.5. Direct comparison between sensitivity and visual preference

Results from sensitivity and visual preference tasks were compared. For this, data from both experimental task conditions were converted to z-scores to ensure they had the same units. Data from the visual preference task was arcsine transformed prior to being z-scored transformed to stabilise variance (McDonald, 2014). After the data was standardised, a 2 × 3 × 5 repeated measures ANOVA was conducted where the main effects included *Task* (sensitivity, visual preference), *SS* (0.25, 1.25, 2.25), and *TS* (0.25, 0.75, 1.25, 1.75, 2.25). Due to z-score transformation of both data sets, the main effect of *Task* is not meaningful since the mean and standard deviation across both tasks is 0 and 1. The present analysis aims to compare responses between visual preference and sensitivity, and as such here we report *interactions* involving *Task*

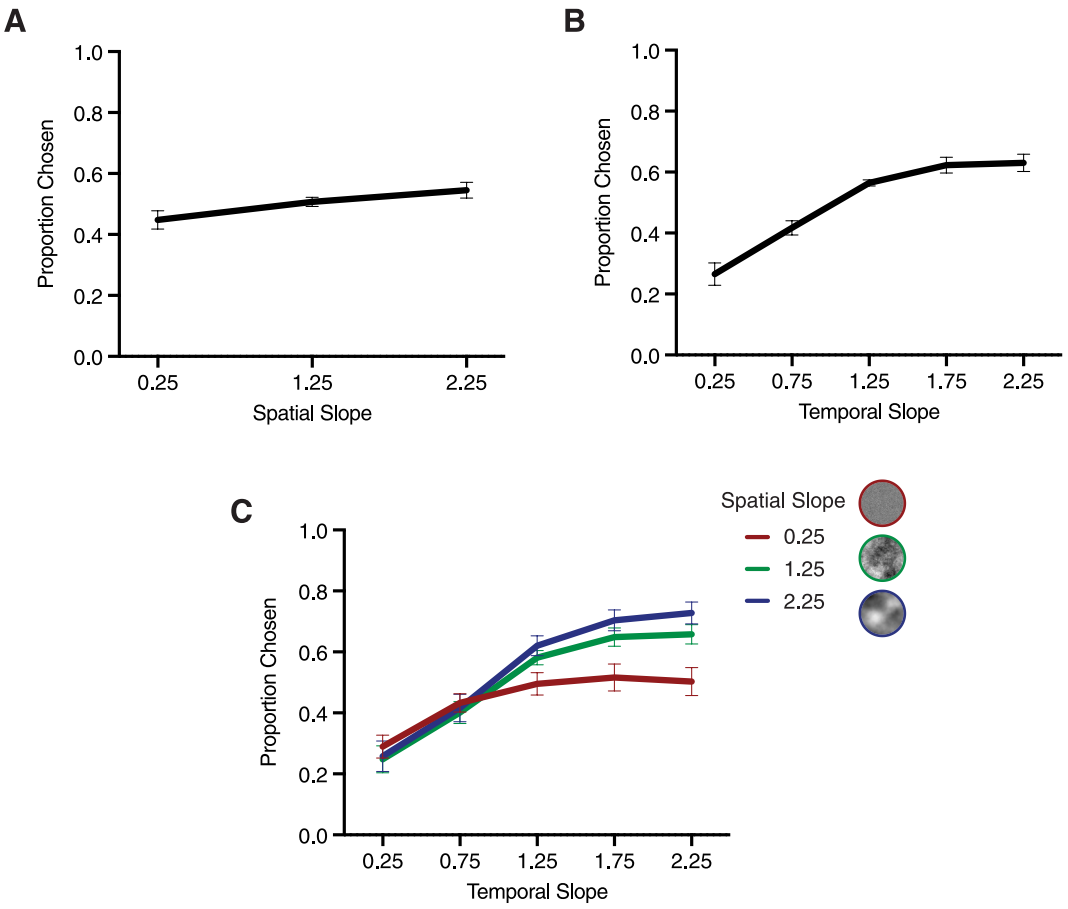


**Fig. 5.** Main effects and interactions from the visual preference analysis. A) Main effect of SS—proportion chosen is plotted as a function of SS. The values plotted were averaged across all TS conditions. Visual preference was highest for steep SS conditions (e.g. 2.25) B) Main effect of TS—proportion chosen is plotted as a function of TS. The values plotted were averaged across all SS conditions. Visual preference was highest for steep TS conditions (e.g. 2.25). C) Interaction between SS and TS—proportion chosen plotted as a function of TS for each SS condition (red, green, blue lines correspond to SS 0.25, SS 1.25, SS 2.25 respectively). Error bars depict SEM between subjects. See Fig. S-4 in the Supplementary Materials to view the same plots with individual datapoints. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

rather than the main effect of *Task* itself ($F_{1,30} = 0.744$, $p = 0.395$, $\eta_p^2 = 0.024$). All main effects and interactions from this analysis are summarised in Table S-3 in the Supplementary Materials.

Interactions

The two-way interactions between *Task* × *SS* ($F_{1.560,46.789} = 5.248$, $p = 0.014$, $\eta_p^2 = 0.149$) and between *Task* × *TS* ($F_{1.883,56.487} = 9.693$, $p < 0.001$, $\eta_p^2 = 0.244$) were significant, and the direction of these interactions is depicted in Fig. 6A and B. These effect sizes are considered large as per Cohen's (1988) conventions. The shape of the response profiles between task conditions as a function of SS (Fig. 6A) is dissimilar, where an inverted U-shape peaking at SS 1.25 is observed for the sensitivity task while a linear increase is observed between SS 0.25 to SS 2.25 for the visual preference task.

The three-way interaction *Task* × *SS* × *TS* ($F_{4.625,138.749} = 6.375$, $p < 0.001$, $\eta_p^2 = 0.175$) was also significant, and the effect size can be considered large (Cohen, 1988). The direction of this interaction is depicted in Fig. 6C. The same pattern of results for visual preference as described above is observed for all SS conditions—a monotonic increase from *shallow* to *steep* TS conditions. Conversely, sensitivity resembles an inverted U-shaped pattern of results with peak responses that systematically change depending on the SS—i.e. a rightward shift in peak sensitivity as a function of TS with increasing SS. The similarity between task conditions increases from SS 0.25 to SS 2.25 (i.e. similarity in shape and distance between solid and dashed lines in Fig. 6C), suggesting that visual preference and sensitivity are more concordant for SS 1.25 and SS 2.25.

### 2.2.6. Exploratory Factor Analysis

We also conducted an Exploratory Factor Analysis to assess how well a small set of variables can capture common variance in our sensitivity and preference datasets (Child, 1990; Suhr, 2005). We also conducted this analysis to assess how similar the factors extracted from both datasets would be. Our exploratory factor analysis conducted across all conditions (15 total – 3 SS, 5 TS) yielded three factors for both datasets, and these factors explained 72.3% of the variance in the case of sensitivity, and 85.2% in the case of preference. These findings suggest that participants were performing in a relatively systematic way during the preference task compared to the sensitivity task—despite not having a concrete definition of how visual preference should be judged. If this was not the case, the factors pulled from the EFA would not be able to explain a large degree of variance in the preference dataset. To view a plot of the loadings of the EFA factors for each dataset, see Fig. 7 (and to compare to the raw sensitivity and preference data see Figs. 4 and 5 respectively).

The factors across datasets differ. For sensitivity, three factors appear to correspond to the pattern of sensitivity across the different SS conditions. Factor 1 (36.7% explained variance) appears to correspond to the sensitivity profile of SS 2.25; Factor 2 (23.7% explained variance) the sensitivity profile of SS 0.25 and SS 1.25, and Factor 3 (11.9% explained variance) a scaling factor to account for different magnitudes of the response profiles to SS 0.25 and 1.25, and also SS 2.25 to a degree

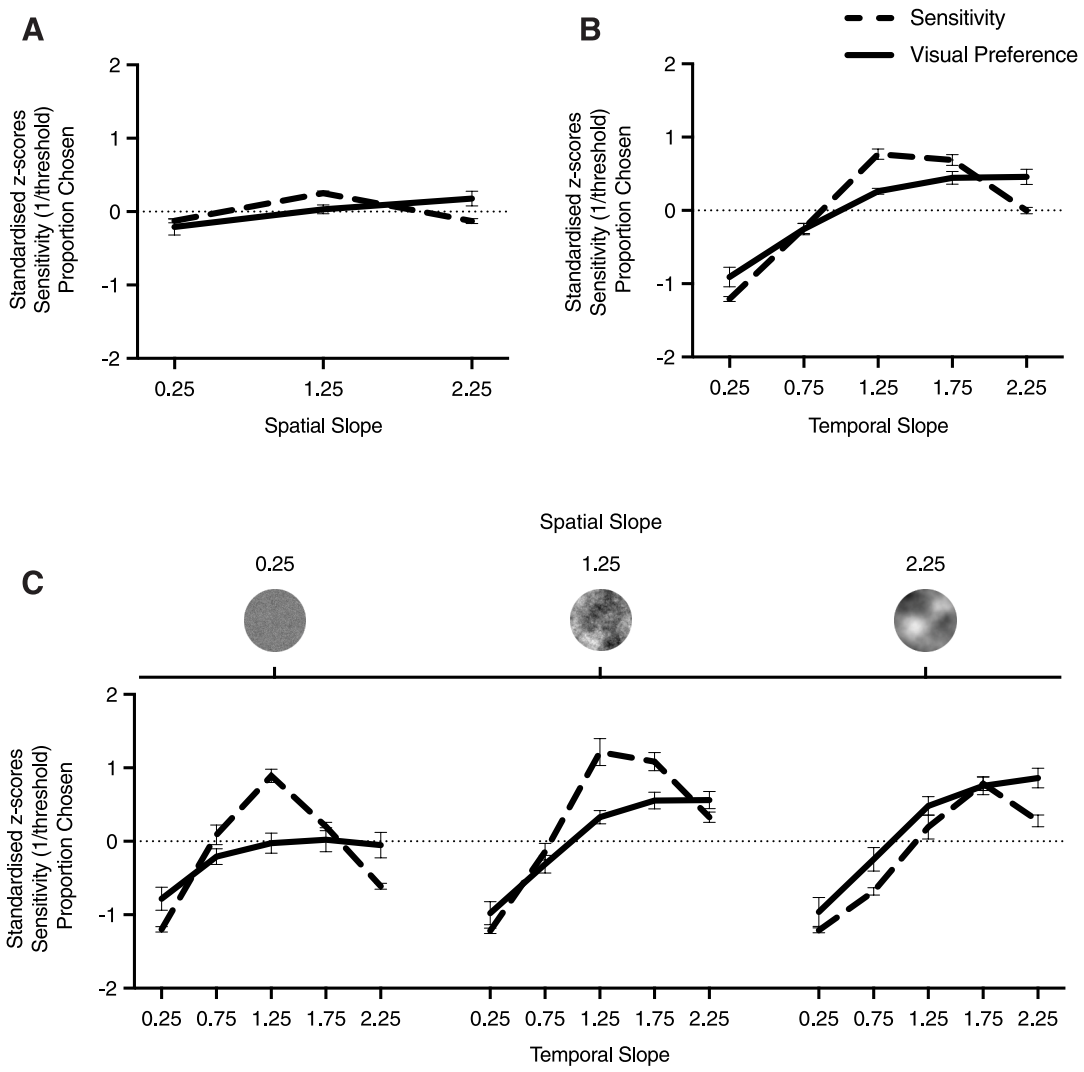


**Fig. 6.** Measuring the degree of concordance between sensitivity and visual preference. A) Comparison between sensitivity and visual preference as a function of SS averaged across TS conditions. A discrepancy is observed between tasks where sensitivity resembles an inverted U-shape peaking for SS 1.25 while visual preference increases systematically from shallow to steep (left to right) TS conditions. B) Comparison between visual preference and sensitivity as a function of TS averaged across SS conditions. The response curves between the two task conditions appear similar. However, at TS 2.25 visual preference increases, while sensitivity drops C) Comparison between sensitivity and visual preference measures across SS conditions as a function of TS (left to right columns SS 0.25, SS 1.25, SS 2.25). Error bars depict SEM between subjects.

(scaling responses down for TS 2.25). For preference, the three factors also appear to correspond to the pattern of preference across the different SS conditions. Factor 1 (45.6% explained variance) appears to correspond to the preference profile for SS 0.25, Factor 2 (26.3% explained variance) the preference profile for SS 2.25, and Factor 3 (11.3% explained variance) the preference profile for SS 1.25. See Tables S-4 and S-5 in the Supplementary Materials for a summary of the factors for sensitivity and visual preference respectively.

### 2.3. Discussion

In *Experiment 1*, we sought to characterise sensitivity and visual preference toward stimuli which differed in their $1/f^{\alpha}$ spatiotemporal spectra. Following research conducted in the *spatial* domain (Spehar et al., 2003, 2015, 2016; Juricevic et al., 2010), we predicted that sensitivity and preference would overlap across all of our stimuli and peak for the most *natural* stimulus in our set. We confirmed that the most *natural* stimulus in our set elicited the highest sensitivity (i.e. it was the easiest stimulus to discriminate)—however, it was *not* the most preferred on average. We find that a stimulus with a spatiotemporal spectrum beyond what is found *on average* in nature—SS 2.25 TS 2.25—was the most preferred. Following this finding, the question arises: *Why does this discrepancy occur?* Before addressing this, we will first discuss the pattern of results observed for *sensitivity* and *visual preference* separately.

#### 2.3.1. Sensitivity

While we observed maximum discrimination sensitivity toward the most natural stimulus in the set—SS 1.25 TS 1.25—we also observed systematic differences in sensitivity dependent on both SS and TS. When sensitivity is plotted as a function of TS, the response profile resembles an inverted U-shape with a peak that shifts rightwards from *shallow* to *steep* SS conditions (Fig. 4C). This essentially corresponded to heightened sensitivity for stimuli that had $1/f^{\alpha}$ spectra close in space and time. These findings are in accordance with Billock et al. (2001a), who also found lower JND thresholds for stimuli with corresponding spatiotemporal spectra.

Based on the spatiotemporal contrast sensitivity function (CSF), one may in fact predict the opposite of these findings—sensitivity across SFs becoming progressively more band-pass as TF increases (Kelly, 1979). If

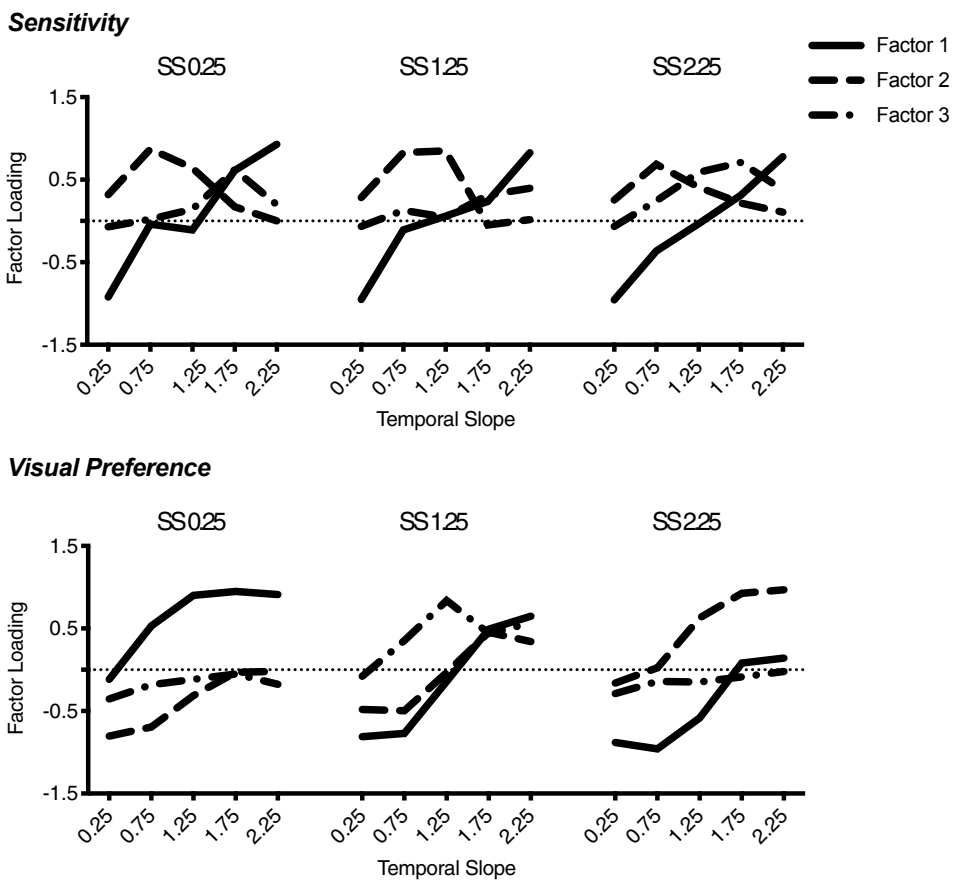


**Fig. 7.** Plots of the EFA factor loadings for sensitivity (top panel) and preference (bottom panel) datasets. To compare to the raw sensitivity and preference data see Figs. 4 and 5 respectively. For sensitivity, three factors appear to correspond to the pattern of sensitivity across the different SS conditions. Factor 1 (solid line): corresponds to the sensitivity profile of SS 2.25; Factor 2 (dashed line): corresponds to the sensitivity profile of SS 0.25 and SS 1.25; Factor 3 (dashed dotted line): scaling factor to account for different magnitudes of the response profiles to SS 0.25 and 1.25, and also SS 2.25 to a degree (scaling responses down for TS 2.25). For preference, the three factors also appear to correspond to the pattern of preference across the different SS conditions. Factor 1: preference profile for SS 0.25; Factor 2: preference profile for SS 2.25; Factor 3: preference profile for SS 1.25.

this was the case, in the present study we should have observed relatively equal sensitivity across SS conditions when the corresponding TS was *shallow* (i.e. rapid luminance modulation, TS 0.25 and TS 0.75). However, given that the spatiotemporal CSF is based on *narrowband* measures of spatial and temporal frequency, the pattern of sensitivity it describes may not generalise to the *broadband* stimuli used in the present study nor others (Simoncini et al., 2012; Gekas et al., 2017; Vacher et al., 2018). *So, what can account for our pattern of results?* While we cannot offer a detailed model of how the visual system is tuned to *broadband* distributions of spatial and temporal frequencies, we can speculate as to why we observe heightened sensitivity for stimuli with $1/f^{\alpha}$ spectra that are close in space and time.

While natural scenes on *average* have spatiotemporal $1/f^{\alpha}$ spectra with an α close to 1 (Billock et al., 2001b; Dong & Atick, 2009)—this averaged α value does not capture the *spatiotemporal* correlations that may exist across certain natural scenes. For instance, how is the movement of a swarm of bees captured by this measure? Or the movement of clouds? Dong and Atick (2009) investigated this by measuring spatiotemporal correlations in amplitude spectra across a wide range of movies of natural scenes. The spatial and temporal spectra of natural scenes was found to generally be *non-separable*, suggesting that these components are typically concordant. As such, it is possible that the visual system is sensitive to *spatiotemporal* concordance in $1/f^{\alpha}$ spectra—consistent with the observed heightened sensitivity for stimuli with $1/f^{\alpha}$ spectra close in both space and time.

### 2.3.2. Visual Preference

Contrary to our hypotheses, instead of observing peak preference for the most *natural* stimulus in the set, we observe peak preference for the *steepest* slope in both space and time—2.25. The most preferred SS was 2.25 followed by 1.25 then 0.25 (Fig. 5A), and the same systematic increase in preference was observed from *shallow* to *steep* TS conditions—albeit at different overall magnitudes depending on the SS condition (Fig. 5C).

It has been well established across a multitude of studies using static images that the most preferred SS is ~1.25 (Spehar et al., 2003, 2015; Juricevic et al., 2010; Spehar & Taylor, 2013). Our results suggest that the addition of a *temporal* component to an otherwise *static* stimulus can profoundly change its visual appeal. When investigating the emotional effects of dynamic textures, Toet et al. (2011) presented participants a variety of natural scene movies and found that the *speed* of the stimulus most highly correlated with ratings of *relaxation* and *pleasure*—the slower the stimulus, the higher its relaxation and pleasure ratings. In addition, *spatial regularity*—i.e. the degree in spatial variation within a stimulus—positively correlated with measures of pleasure and relaxation.

As such, stimuli with *slow* temporal dynamics and *high* spatial regularity could be considered the most preferred in Toet et al. (2011)—analogous to our findings that the stimulus with low spatial variation that slowly varied in luminance over time was most preferred (SS 2.25 TS 2.25). Indeed, the appearance of a random noise stimulus with *steep* $1/f^{\alpha}$ spectra in space and time has been described by a number of participants as containing "rather coherent large masses undulating very slowly" (Billock et al., 2001a), corresponding to the factors described by Toet et al. (2011). Visual preference toward the spatiotemporal $1/f^{\alpha}$ amplitude spectrum appears therefore to be driven mostly by the *temporal* component of our stimuli (the *slower* the *better*) and to a lesser extent its *spatial* component (the *smoother* the *better*).

### 2.3.3. The relationship between sensitivity and visual preference

The divergence we observe between sensitivity and visual preference is contrary to our initial expectation that the *exact* same pattern of results would be observed across *spatial* and *temporal* slope combinations, following findings in the *spatial* domain (Spehar et al., 2015). As stated

above, the addition of a temporal component to an otherwise static stimulus influences its visual appeal. On the other hand, sensitivity toward these stimuli seems to accord with the spatiotemporal correlations that exist across natural scenes. *So why do we observe this divergence between sensitivity and preference in the temporal domain?*

It is possible that our preference findings are consistent with a Bayesian prior for "slow" and "smooth" moving objects. In their investigation of how motion illusions arise, Weiss et al. (2002) sought to model how motion is estimated by the visual system. Their approach is based on two plausible assumptions: 1) local measurements made by the visual system are inherently noisy and 2) that velocities in the natural world tend to be slow—i.e. most objects in nature are stationary, or if they move, they do so slowly. The model was able to accurately predict a range of psychophysical data related to perceived direction and speed (e.g. gratings across a range of contrast conditions, a translating rhombus, and plaids). Weiss et al. (2002) concluded that ideal observers assume that velocities are 'slow and smooth'—somewhat analogous to the most preferred stimulus in our set (SS 2.25, TS 2.25)—and that this bias stems from what exists in our natural environment.

It may also be possible that the divergence observed in the *temporal* domain may be reflective of a different judgement criterion being used in the preference task. It seems sensible for the visual system to be sensitive to spatiotemporal correlations that most commonly exist in nature, particularly $\alpha \approx 1$ (Billock et al., 2001b; Dong and Atick, 2009). However, preference may be based on dynamic contextual cues that may signal preferred environmental conditions. Take for instance the scene of a beach on a calm versus stormy day. Theoretically, both scenes would have similar *spatial* amplitude spectra. However, their *temporal* amplitude spectra would differ considerably—one having a *steep* spectrum, and the other *shallow*. As such, the *temporal* amplitude spectrum may provide additional information that can be used as an environmental cue. As to why the addition of a temporal component corresponds with heightened preference for stimuli with increased *spatial regularity*, e.g. SS 2.25, we cannot provide a complete account. However, whilst speculative, perhaps this change in preference may be related to higher level preference judgments of environmental cues—e.g. a *smooth* vs. a *spiky* surface (Viengkham et al., 2019).

We must acknowledge, however, that while we speculate our preference findings may be indicative of preferred environmental conditions—the stimuli we used in the present study were random noise textures. Further work using movies of actual natural scenes with analogous variations in their spatiotemporal spectra will need to be compared to stimuli used in our study. In addition, further work directly probing the criteria our subjects used to infer their preference judgments with synthetic noise stimuli may reveal whether our notions hold.

#### 2.3.4. Limitations

The divergence between sensitivity and visual preference is apparent for *steeper* TS conditions. However, it is possible that differences in the way the data were collected across sensitivity and visual preference tasks may account for the observed differences. In the visual preference task, *all* spatial and temporal slope combinations were compared—providing a *spatiotemporal* measurement of visual preference across slope conditions. In contrast, sensitivity was separately measured for each spatial and temporal slope condition, resulting in the sensitivity task containing 15 blocks total (3 SS × 5 TS). SS was fixed within each block, and TS varied across trials based on an adaptive staircase procedure that modelled possible increments and decrements around one specified TS condition. As such, it may be possible that visual preference differs when measured in a way such that the SS or TS is *fixed* within a block in comparison to when preference is measured in an *intermixed* fashion. To address this, we investigated whether the visual preference of our stimuli changes when SS is fixed within each block in Experiment 2.

## 3. Experiment 2: The effect of presentation order on visual preference

We conducted a follow up experiment to investigate whether or not the visual preference of our stimuli changes when SS is fixed within each block (*spatial fixed*). We directly compared the preference data from the present study to the visual preference data collected in *Experiment 1* (*intermixed* presentation). For brevity, the *Methods* section, the analysis of the raw data (prior to z-score transformation), as well as a comparison between the preference data collected here and sensitivity collected in *Experiment 1*, are included in the Supplementary Materials.

### 3.1. Results

Prior to analysis, the data across both experiments were z-score transformed such that the corresponding mean and standard deviation of each SS condition became 0 and 1 to ensure both task conditions had the same units to allow a direct comparison between them. Visual preference data were arcsine transformed prior to being z-score transformed as a variance stabilisation measure (McDonald, 2014). All the raw data from the present experiment has been anonymised and is available on the following OSF project webpage: https://osf.io/avzhc/.

#### 3.1.1. Comparing visual preference across spatial fixed and intermixed presentation orders

The data were analysed using a 2 × (3 × 5) mixed ANOVA with *Presentation Type* as the between-subject factor, and *SS* and *TS* as within-subject factors. Due to z-score transformation, the direct comparison between presentation types is not meaningful since the mean and standard deviation for both presentation types is equal to 0 and 1. Given the present analysis compares the pattern of results between fixed and intermixed presentation types, here we focus on interactions rather than the main effect of *Presentation Type* itself ($F_{1,60} = 1.000$, $p = 0.321$, $\eta_p^2 = 0.016$). All main effects and interactions from this analysis are summarised in Table S-7 in the Supplementary Materials.

#### 3.1.2. Interactions: Spatial Fixed vs. Intermixed comparison

The two-way interaction *Presentation Type* × *Temporal Slope* ($F_{1.286,77.147} = 2.811$, $p = 0.675$, $\eta_p^2 = 0.016$) and the three-way interaction *Presentation Type* × *Spatial Slope* × *Temporal Slope* ($F_{5.232,313.938} = 1.459$, $p = 0.200$, $\eta_p^2 = 0.024$) were not significant. The interaction *Presentation Type* × *Spatial Slope* ($F_{1,60} = 1.000$, $p = 0.321$, $\eta_p^2 = 0.016$) was also not significant. However, the comparison between these two main effects is not meaningful due to the way the data was z-score transformed. The direction of these interactions is depicted in Fig. 8. All effect sizes are considered small (Cohen, 1988). The only observable difference in the pattern of results between presentation conditions is for SS 0.25, where preference distinctly peaks for TS 1.25 when the presentation type is *fixed*, while preference plateaus between TS 1.25 and TS 2.25 when the presentation type is *intermixed*.

#### 3.1.3. Discussion

We found no major differences in the pattern of results between *fixed* and *intermixed* presentation conditions. The most preferred stimulus across presentation conditions was still the *coarsest* and *slowest* stimulus in the set (SS 2.25 and TS 2.25), replicating the findings of *Experiment 1*.

## 4. Experiment 3: Measuring sensitivity and visual preference to temporal $1/f^{\alpha}$ amplitude spectra ranging between 0.50 and 2.50 (step size 0.5)

Following the findings of *Experiments 1* and *2*, we explored whether the same pattern of results for sensitivity and visual preference would be observed for a differing range of TS conditions. With a finer sampling rate, we can better infer how narrowband or broadband the tuning we

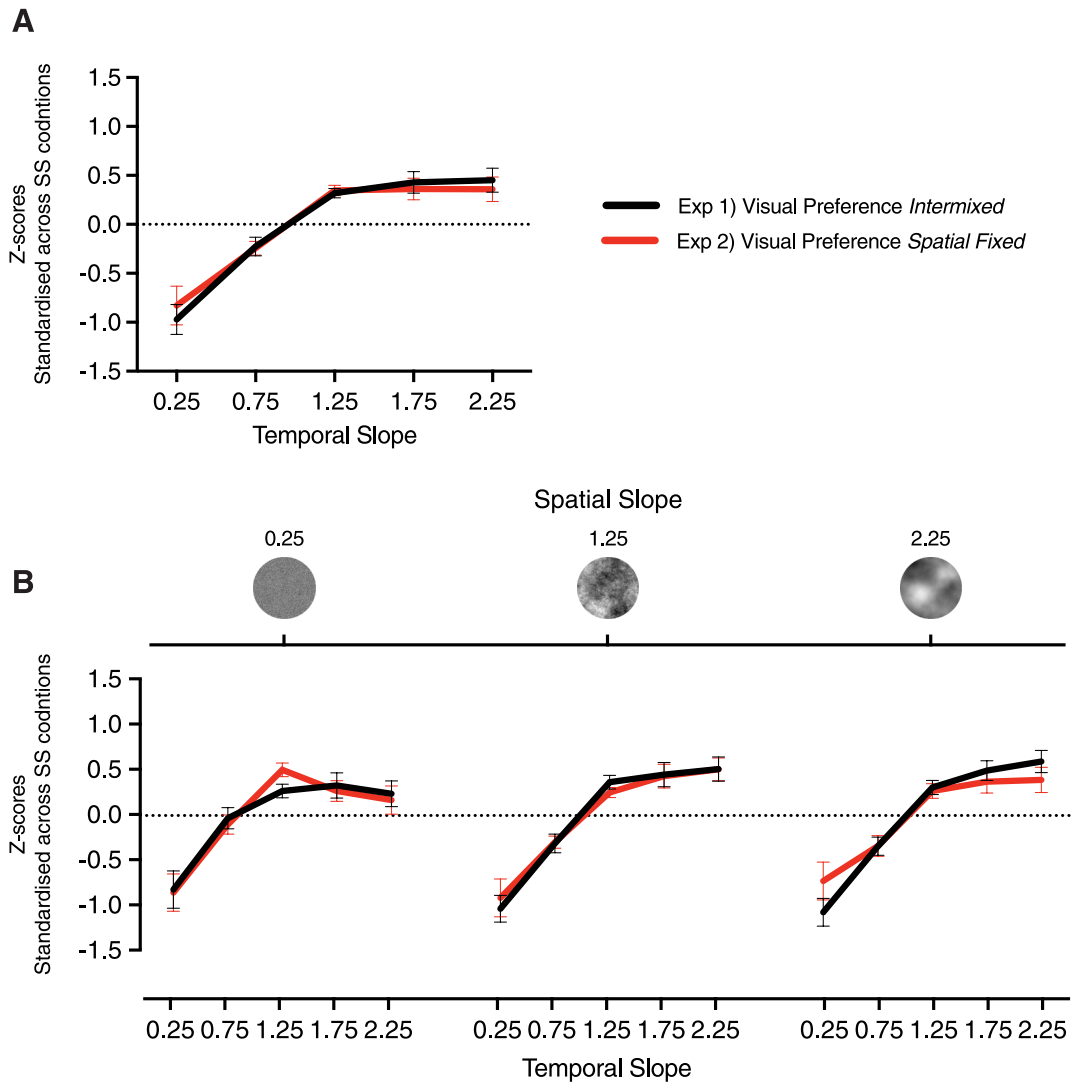


**Fig. 8.** Comparison of visual preference between intermixed (black lines) and spatial fixed (red lines) presentation orders. A) Preference averaged across SS conditions plotted as a function of TS. The pattern of results across presentation conditions overlap where a systematic increase in preference from shallow to steep TS conditions is observed. B) Preference plotted as a function of TS for each SS condition separately. Irrespective of SS condition, the same increase in preference as depicted in (A) is observed for each presentation type. For SS 0.25, however, there appears to be a slight difference in peak preference, where TS 1.25 is more preferred when presented in a fixed manner. Error bars depict SEM between subjects. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

observe in our sensitivity experiment is (*Experiment 1*), as well as whether visual preference continues to increase beyond TS 2.25 (i.e. *is there a point where the stimulus appears too slow and is less preferred?*) We conducted the same sensitivity and visual preference experiments in *Experiment 1* using the same SS conditions ($\alpha = 0.25, 1.25, 2.25$), but a different range of TS conditions ($\alpha = 0.50, 1.00, 1.50, 2.00, 2.50$). Given the extensive analyses conducted in our previous experiments, we did not focus here on the direct comparison between sensitivity and visual preference. Here, we describe the shape of these tuning curves when the data from *Experiment 1* (TS conditions: 0.25, 0.75, 1.25, 1.75, 2.25) are combined with the data from the present study after z-score transformation. For brevity, the *Methods* section and the analysis of the raw data (prior to z-score transformation) are included in the Supplementary Materials.

### 4.1. Results

Data across both *Experiment 1* and *3* were transformed into z-scores to allow a direct comparison between them. Visual preference data were arcsine transformed prior to being z-score transformed as a variance stabilisation measure (McDonald, 2014). All raw data (sensitivity and visual preference) has been anonymised and is available on the following OSF project webpage: https://osf.io/dfkj7/.

#### 4.1.1. Comparison between Experiment 1 and Experiment 3

Given the design of both experiments (different samples, different set of TS conditions in each sample), we are unable to conduct a statistical analysis (e.g., an ANOVA) on the combined data. Instead, here describe the shape of tuning curves for both experiments, and fit the data using polynomials (1st, 2nd, and 3rd degree) determined by Akaike Information Criterion (AIC) values. These values are calculated based on the trade-off between goodness of fit of a model and its complexity by estimating the amount of information lost (Akaike, 1974). AIC values are used to indicate which model was best suited to explain the data, whereby the model with the lowest AIC value (least information lost) amongst other models tested indicates it is the best fit. For a summary of AIC values across polynomials, refer to Table S-13 (sensitivity) and Table S-14 (visual preference) in the Supplementary Materials.

#### 4.1.2. Sensitivity

When sensitivity is averaged across all TS conditions, we observe the pattern of tuning as a function of SS (Fig. 9A), which resembles an inverted U-shaped function peaking at SS 1.25 (2nd degree polynomial, $R^2 = 0.89$). This fit, however, was conducted using only three points—so the data in this case may be overfitted. When sensitivity is averaged across all SS conditions, we can observe the pattern of tuning as a function of TS (Fig. 9B), which also resembles an inverted U-shaped function peaking ~ TS 1.50 (2nd degree polynomial, $R^2 = 0.90$). When plotting all the data as a function of TS, different tuning functions can be observed for each SS separately (Fig. 9C). Here we can observe that stimuli within the *natural range* (SS 1.25 and TS 1.00 to 1.75) elicit the highest sensitivity in the stimulus set (Fig. 9C, green line). While sensitivity to SS 0.25 and 2.25 are generally lower overall as a function of TS (Fig. 9C, red line and blue line respectively), we observe heightened sensitivity for stimuli which are concordant in space and time.

#### 4.1.3. Visual preference

When averaged across all TS conditions, the visual preference tuning pattern as a function of SS is essentially a flat line with a *slight* increase in preference from SS 0.25 to SS 2.25 (1st degree polynomial, $R^2 = 0.46$) (Fig. 10A). Across all SS conditions, the pattern of tuning for visual preference as a function of TS resembles a logarithmic function starting to asymptote ~ TS 1.75 (2nd degree polynomial, $R^2 = 0.90$) (Fig. 10B). When plotting all the data as a function of TS, tuning functions can be observed for each SS separately (Fig. 10C). Here we observe that preference increases as a function of TS, where SS 2.25 is *slightly* more preferred than SS 1.25, followed by SS 0.25. However, for TS conditions 0.25 to 1.00, SS 0.25 is most preferred. The shape of these tuning curves is similar across SS conditions. For SS 1.25 and SS 2.25, the response curves are essentially the same and resemble logarithmic functions starting to asymptote ~ TS 1.75 (SS 1.25: 2nd degree polynomial, $R^2 = 0.91$; SS 2.25: 2nd degree polynomial, $R^2 = 0.83$). For SS 0.25, the tuning curve somewhat resembles a logarithmic function like SS 1.25 and SS 2.25 (2nd degree polynomial, $R^2 = 0.47$), however the increase in preference is shallower between TS 0.25 and TS 1.50, and preference appears to decrease between TS 2.00 and TS 2.50.

#### 4.1.4. Discussion

The present study has shown that we observe a similar pattern of sensitivity and visual preference tuning toward the *spatiotemporal* $1/f^{\alpha}$ amplitude spectrum across a new range of TS conditions ($\alpha = 0.50, 1.00, 1.50, 2.00, 2.50$). By combining the results from *Experiment 1* and *Experiment 3*, we were able to describe and fit response curves across sensitivity and visual preference measures. We find that the fit of these response curves differs substantially between these two measures, further indicating that a divergence in sensitivity and preference occurs in the *temporal* domain. A novel finding of the present study is that visual preference starts to decrease for SS 0.25 when its corresponding TS is *too steep* (TS 2.50). It would be interesting in future research to test at what point (if any) visual preference starts to decrease for SS 1.25 and SS 2.25 at *steeper* and *steeper* TS conditions (e.g. TS 3.00 or 4.00). Future research could also include a *stationary* static condition (SS manipulations *only*) to be able to directly compare the influence of *motion* in visual preference.

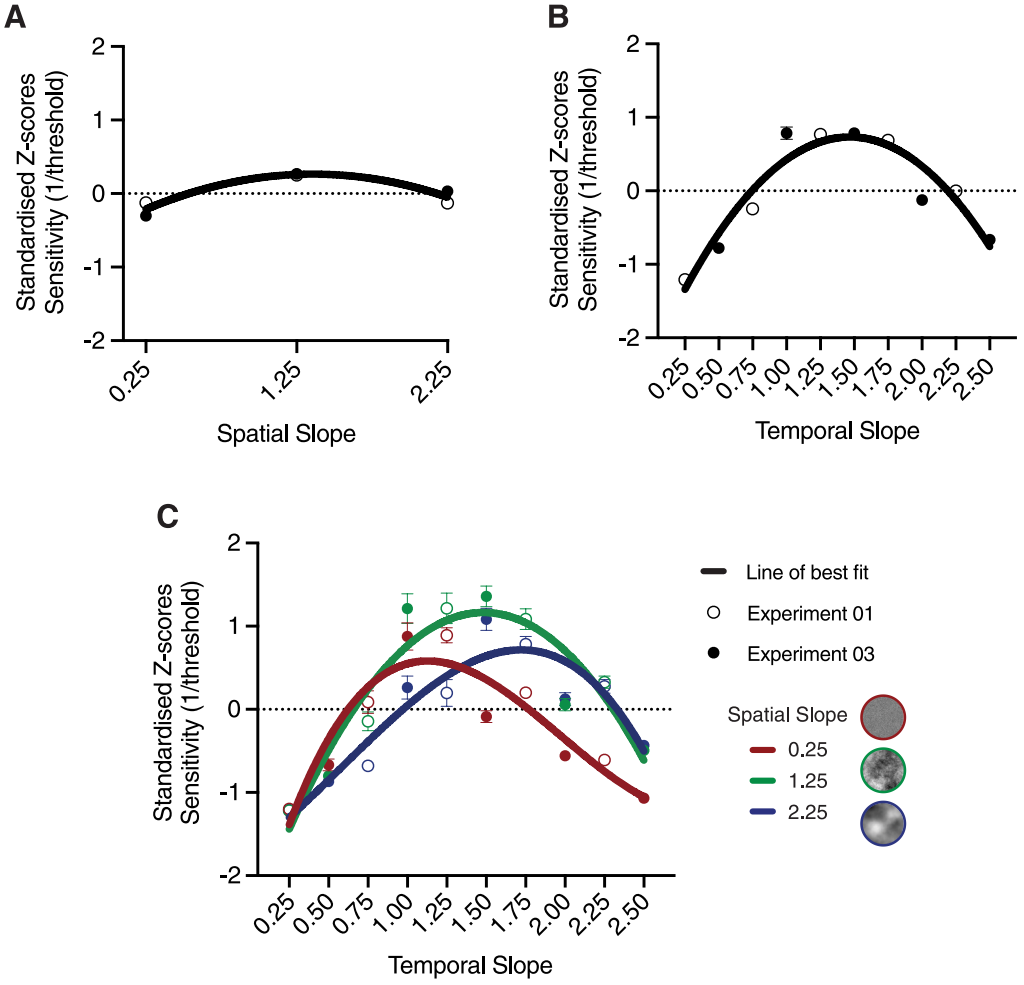


**Fig. 9.** Plotting discrimination sensitivity across Experiment 1 and Experiment 3. A) Main effect of SS—sensitivity plotted as a function of TS. The values plotted were averaged across all TS conditions within each experiment separately. An inverted U-shaped pattern peaking for SS 1.25 is observed. B) Main effect of TS—sensitivity plotted as a function of TS. The values plotted were averaged across all SS conditions. An inverted U-shaped pattern peaking around TS 1.50 is observed. C) Interaction between SS and TS—sensitivity plotted as a function of TS for each SS condition separately (red, green, blue lines correspond to SS 0.25, SS 1.25, SS 2.25 respectively). Each SS condition has a slightly different pattern of results. Overall, the most natural stimuli in the set elicit the highest sensitivity (e.g. SS 1.25 TS 1.50). Note: Error bars indicate the SEM. The polynomial degree chosen for each fit was based on the AIC calculated across 1st, 2nd, and 3rd degree polynomial fits (see Table S-13 in the Supplementary Materials for a summary). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

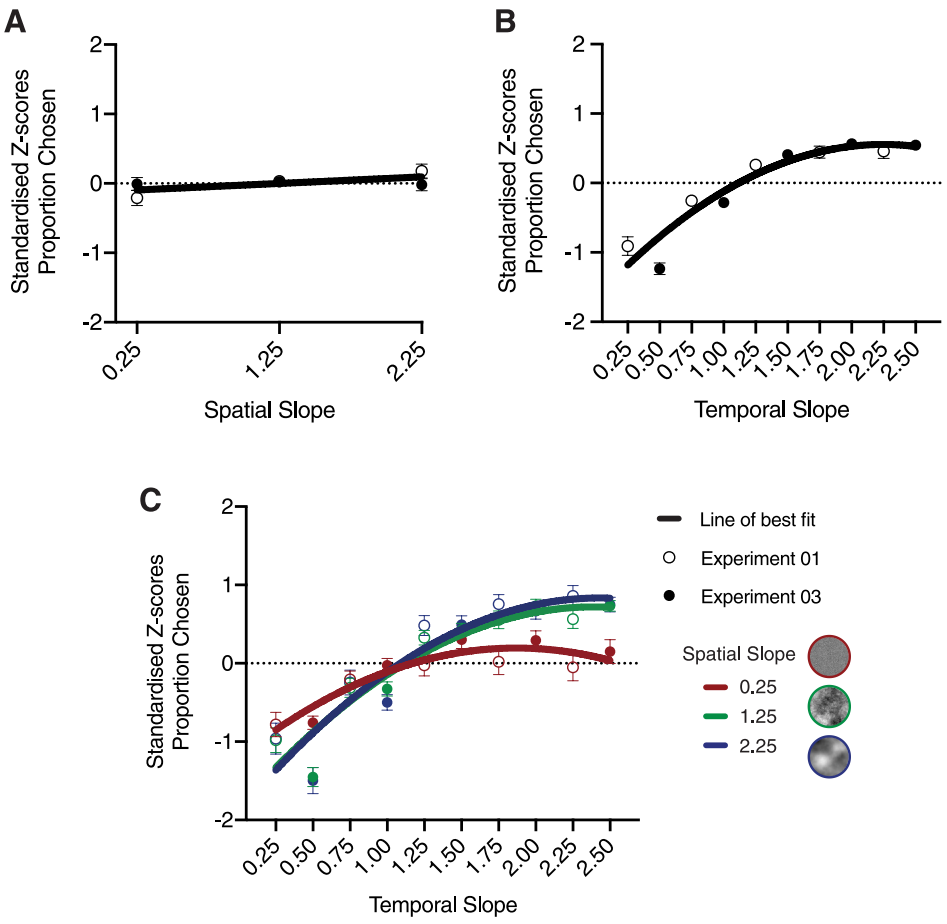


**Fig. 10.** Plotting visual preference across Experiment 1 and Experiment 3. A) Main effect of SS—proportion chosen is plotted as a function of SS. The values plotted were averaged across the TS conditions within each experiment separately. Visual preference increased linearly from SS 0.25 to SS 2.25, however this increase is minimal. B) Main effect of TS—proportion chosen is plotted as a function of TS. The values plotted were averaged across all SS conditions. Visual preference was highest for steep TS conditions (e.g. TS 2.00 to 2.50). C) Interaction between SS and TS—proportion chosen plotted as a function of TS for each SS condition (red, green, blue lines correspond to SS 0.25, SS 1.25, SS 2.25 respectively). Across all SS conditions, preference increased as a function of increasing TS from 0.25 to 2.50. Note: Error bars indicate the SEM. The polynomial degree chosen for each fit was based on the AIC calculated across 1st, 2nd, and 3rd degree polynomial fits (see Table S-14 in the Supplementary Materials for a summary). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## 5. General Discussion

The series of experiments described here are some of the first to systematically measure and compare sensitivity and preference toward the $1/f^{\alpha}$ amplitude spectrum in both space *and* time. Following work conducted purely on the *spatial* domain (Spehar et al., 2015), we predicted that the most natural stimulus both *spatially* and *temporally* would be most preferred and the easiest to discriminate. We also predicted a strong concordance between preference and sensitivity across all other spatial and temporal variations in $1/f^{\alpha}$ spectra.

Contrary to our hypotheses, preference for natural $1/f^{\alpha}$ spectra ($\alpha \approx 1$) only holds for static stimuli. The most preferred stimulus in our set was found to have *steep* $1/f^{\alpha}$ spectra in space and time, corresponding to the *coarsest* and *slowest* stimulus in the set (SS 2.25 TS 2.25). In addition, the *slowest* stimulus was always preferred irrespective of its *spatial* properties (TS 2.25). Our predictions held for sensitivity as the most discriminable stimulus were the most *natural* in the set (SS 1.25 TS 1.25). Interestingly, peak sensitivity shifted depending on the *spatial* properties of a stimulus, such that stimuli with $1/f^{\alpha}$ spectra concordant in space and time were easier to discriminate (SS/TS—*shallow/shallow*, *intermediate/intermediate*, *steep/steep*). This pattern of results corroborates the findings of Billock et al. (2001a), suggesting the visual system is tuned to $1/f^{\alpha}$ spectra concordant in space and time—a property which typically exists across natural scenes (e.g. small insects typically move quickly, big clouds typically move slowly) (Dong & Atick, 2009).

It seems sensible for the visual system to be maximally tuned to $1/f^{\alpha}$ spectra that exist *commonly* across natural scenes—however, this does not mean one necessarily *prefers* such spectra. While we do observe a strong concordance between sensitivity and preference to a certain point the divergence between measures for *steep* 1/f slope conditions cannot be dismissed. We speculate that with the addition of a temporal component, the preference of a stimulus may be based on either a Bayesian prior to 'slow and smooth' velocities (Weiss et al., 2002), or perhaps dynamic contextual cues that signal preferred environmental conditions. As mentioned previously, in comparisons between a beach on a calm day vs. a stormy day—the spatial $1/f^{\alpha}$ spectra across these comparisons are relatively similar. The only spectral information discerning both comparisons is in the *temporal* domain, and in general scenes which are more preferred move *slowly* (Toet et al., 2011). However, it is important to note that the stimuli we use in our experiments are synthetic filtered textures. As such, the extension of our findings to real-world scenes is at present speculative and needs further investigation.

Future work measuring the spatiotemporal $1/f^{\alpha}$ amplitude spectrum of a wider range of natural scenes than done previously (Billock et al., 2001b; Dong & Atick, 2009) may help disentangle the dissociation we observe between sensitivity and visual preference. For example, Toet et al. (2011) had participants rate a wide range of natural movies on a variety of emotional attributes (e.g. relaxation, pleasure, arousal) and spatiotemporal descriptors (e.g. spatial regularity, speed frequency, temporal regularity). They subsequently modelled the extent to which the rated spatiotemporal descriptors of each movie could account for its emotional attributes. Instead of using the descriptors labelled by participants, it would be interesting to model the extent to which the measured spatiotemporal $1/f^{\alpha}$ amplitude spectrum of their stimuli could explain their findings that *slow* stimuli with *high* spatial regularity were most preferred. This may indicate whether it is indeed the case that scenarios such as those described above (e.g. a beach on a calm day) are preferred based on a criterion related their *temporal* $1/f^{\alpha}$ properties rather than their *spatial* $1/f^{\alpha}$ properties.

## CRediT authorship contribution statement

**Zoey J. Isherwood:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Colin W.G. Clifford:** Conceptualization, Methodology, Software, Resources, Writing - review & editing. **Mark M. Schira:** Resources, Writing - review & editing, Supervision. **Michelle M. Roberts:** Formal analysis, Investigation, Writing - review & editing. **Branka Spehar:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

## Acknowledgements

We thank Sophia Kwan for assistance with data collection, and Harriet Boyd Taylor for comments and edits on early iterations of the manuscript. This research was supported by Australian Research Council (ARC) Discovery Project Grant DP170104018 awarded to BS and the Australian Government Research Training Program (RTP) Scholarship awarded to ZJI.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.visres.2021.01.001.

## References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.
Baker, D. H., & Graf, E. W. (2009). Natural images dominate in binocular rivalry. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 5436–5441.
Billock, V. A., Cunningham, D. W., Havig, P. R., & Tsou, B. H. (2001a). Perception of spatiotemporal random fractals: An extension of colorimetric methods to the study of dynamic texture. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 18*, 2404–2413.
Billock, V. A., de Guzman, G. C., & Scott Kelso, J. A. (2001b). Fractal time and 1/f spectra in dynamic images and human vision. *Physica D: Nonlinear Phenomena, 148*, 136–146.
Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics, 26*, 157–170.
Cass, J., Alais, D., Spehar, B., & Bex, P. J. (2009). Temporal whitening: Transient noise perceptually equalizes the 1/f temporal amplitude spectrum. *Journal of Vision, 9*(12), 11–19.
Child, D. (1990). *The essentials of factor analysis*. Cassell Educational.
Cohen, J. P. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York, NY: Routledge Academic.
Cohn, J. (1894). Experimentelle Unterschungen uber die Gefuhls-betonung der Farben, Helligkeiten, und iherer Combinationen. *Philosphische Studien, 10*, 562–603.
Dong, D. W., & Atick, J. J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems, 6*, 345–358.
Dong, D. W., & Atick, J. J. (2009). Statistics of natural time-varying images. *Network: Computation in Neural Systems, 6*, 345–358.
Eckert, M. P., & Buchsbaum, G. (1993). Efficient coding of natural time varying images in the early visual system. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 339*, 385–395.
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: Optics, Image Science, and Vision, 4*, 2379–2394.
Flitcroft, D. I., Harb, E. N., & Wildsoet, C. F. (2020). The spatial frequency content of urban and indoor environments as a potential risk factor for myopia development. *Investigative Ophthalmology & Visual Science, 61*, 42.
Gekas, N., Meso, A. I., Masson, G. S., & Mamassian, P. (2017). A normalization mechanism for estimating visual motion across speeds and scales. *Current Biology, 27* (1514–1520), Article e1513.
Hansen, B. C., & Hess, R. F. (2006). Discrimination of amplitude spectrum slope in the fovea and parafovea and the local amplitude distributions of natural scene imagery. *Journal of Vision, 6*, 696–711.
Isherwood, Z. J., Schira, M. M., & Spehar, B. (2017). The tuning of human visual cortex to variations in the 1/f(alpha) amplitude spectra and fractal properties of synthetic noise images. *Neuroimage, 146*, 642–657.
Jakel, F., & Wichmann, F. A. (2006). Spatial four-alternative forced-choice method is the preferred psychophysical method for naive observers. *Journal of Vision, 6*, 1307–1322.
Juricevic, I., Land, L., Wilkins, A., & Webster, M. A. (2010). Visual discomfort and natural image statistics. *Perception, 39*, 884–899.
Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69*, 1340–1349.
Kingdom, F. A. A., & Prins, N. (2016). Chapter 5 - Adaptive Methods*. In F. A. A. Kingdom, & N. Prins (Eds.), *Psychophysics* (second ed., pp. 119–148). San Diego: Academic Press.
Knill, D. C., Field, D., & Kersten, D. (1990). Human discrimination of fractal images. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 7*, 1113–1123.
Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39*, 2729–2737.
Kretzmer, E. R. (1952). Statistics of television signals. *Bell System Technical Journal, 31*, 751–763.
McDonald, J. (2014). *Handbook of Biological Statistics* (3rd ed.). Baltimore, Maryland: Sparky House Publishing.
Nguyen, L. Y., & Spehar, B. (2021). Visual adaptation to natural scene statistics and visual preference. *Vision Research, 180*, 87–95.
O'Hare, L., & Hibbard, P. B. (2011). Spatial frequency and visual discomfort. *Vision Research, 51*, 1767–1777.
Olman, C. A., Ugurbil, K., Schrater, P., & Kersten, D. (2004). BOLD fMRI and psychophysical measurements of contrast response to broadband images. *Vision Research, 44*, 669–683.
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*, 35–38.
Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.
Penacchio, O., & Wilkins, A. J. (2015). Visual discomfort and the spatial distribution of Fourier energy. *Vision Research, 108*, 1–7.
Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Pers Soc Psychol Rev, 8*, 364–382.
Simoncini, C., Perrinet, L. U., Montagnini, A., Mamassian, P., & Masson, G. S. (2012). More is not always better: Adaptive gain control explains dissociation between perception and action. *Nature Neuroscience, 15*, 1596–1603.
Spehar, B., Clifford, C. W. G., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers & Graphics, 27*, 813–820.
Spehar, B., & Taylor, R. P. (2013). *Fractals in art and nature: Why do we like them?* (p. 865118) Human Vision and Electronic Imaging XVIII: International Society for Optics and Photonics.
Spehar, B., Walker, N., & Taylor, R. P. (2016). Taxonomy of individual variations in aesthetic responses to fractal patterns. *Frontiers in Human Neuroscience, 10*, 350.
Spehar, B., Wong, S., van de Klundert, S., Lui, J., Clifford, C. W., & Taylor, R. P. (2015). Beauty and the beholder: The role of visual sensitivity in visual preference. *Frontiers in Human Neuroscience, 9*, 514.
Suhr, D. D. (2005). Principal component analysis vs. exploratory factor analysis (paper 203-30). In Proceedings of the thirtieth annual SAS® users group international conference (p. 30).
Tadmor, Y., & Tolhurst, D. J. (1994). Discrimination of changes in the second-order statistics of natural and synthetic images. *Vision Research, 34*, 541–554.
Toet, A., Henselmans, M., Lucassen, M. P., & Gevers, T. (2011). Emotional effects of dynamic textures. *Iperception, 2*, 969–991.
Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics, 12*, 229–232.
Vacher, J., Meso, A. I., Perrinet, L. U., & Peyre, G. (2018). Bayesian modeling of motion perception using dynamical stochastic textures. *Neural Computation, 30*, 3355–3392.
Van Hateren, J. H. (1993). Spatiotemporal contrast sensitivity of early vision. *Vision Research, 33*, 257–267.
Vancleef, K., Read, J. C. A., Herbert, W., Goodship, N., Woodhouse, M., & Serrano-Pedraza, I. (2018). Two choices good, four choices better: For measuring stereoacuity in children, a four-alternative forced-choice paradigm is more efficient than two. *PLoS ONE, 13*, Article e0201366.
Viengkham, C., Isherwood, Z., & Spehar, B. (2019). Fractal-scaling properties as aesthetic primitives in vision and touch. *Axiomathes*.
Webster, M. A., & Miyahara, E. (1997). Contrast adaptation and the spatial structure of natural images. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 14*, 2355–2366.
Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*, 598–604.
Yoonessi, A., & Kingdom, F. A. (2008). Comparison of sensitivity to color changes in natural and phase-scrambled scenes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 25*, 676–684.
Yoshimoto, S., Garcia, J., Jiang, F., Wilkins, A. J., Takeuchi, T., & Webster, M. A. (2017). Visual discomfort and flicker. *Vision Research, 138*, 18–28.
Yoshimoto, S., Jiang, F., Takeuchi, T., Wilkins, A. J., & Webster, M. A. (2019). Adaptation and visual discomfort from flicker. *Vision Research, 160*, 99–107.
Yoshimoto, S., Jiang, F., Takeuchi, T., Wilkins, A. J., & Webster, M. A. (2020). Visual discomfort from flicker: Effects of mean light level and contrast. *Vision Research, 173*, 50–60.