



Published in final edited form as:

*J Vis.* ; 9(10): 12.1–1219. doi:10.1167/9.10.12.

## Temporal whitening: Transient noise perceptually equalizes the 1/f temporal amplitude spectrum

**John Cass,**

School of Psychology, University of Sydney, Sydney, NSW, Australia

**David Alais,**

School of Psychology, University of Sydney, Sydney, NSW, Australia

**Branka Spehar,** and

School of Psychology, University of New South Wales, Sydney, NSW, Australia

**Peter J. Bex**

Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA

### Abstract

Naturally occurring luminance distributions are approximately  $1/f$  in their spatial and temporal amplitude spectra. By systematically varying the spatio-temporal profile of broadband noise stimuli, we demonstrate that humans invariably overestimate the proportion of high spatial and temporal frequency energy. Critically, we find that the strength of this bias is of a magnitude that predicts a perceptually equalized response to the spatio-temporal fall off in the natural amplitude spectrum. This interpretation is supported by our finding that the magnitude of this transient response bias, while evident across a broad range of narrowband spatial frequencies (0.25–8 cycles/deg), decreases above 2 cycles/deg, which itself compensates for the increase in temporal frequency energy previously observed at high spatial frequencies as a consequence of small fixational eye movements (M. Rucci, R. Iovin, M. Poletti, & F. Santini, 2007). Additional temporal masking and adaptation experiments reveal a transiently biased asymmetry. Whereas temporal frequencies  $>4$  Hz mask and adapt 1- and 15-Hz targets, lower masking and adaptation frequencies have much less effect on sensitivity to 15-Hz compared with 1-Hz targets. These results imply that the visual system over-represents its transient input to an extent that predicts an equalized temporal channel response to the low-frequency-biased structure of natural scenes.

### Keywords

spatio-temporal frequency; masking; adaptation; whitening; equalization;  $1/f$ ; natural scenes

### Introduction

Despite their vast physical complexity, natural visual environments possess local luminance distributions that are lawfully correlated across space and time. Pairwise comparisons of intensity levels reveal that more proximal regions of space, spatially or temporally, tend to possess more similar luminosities than more distant regions. This implies that visual scenes are dominated by regions that are more low frequency biased, spatially and temporally, than

---

Corresponding author: John Cass. johnc@psych.usyd.edu.au. Address: School of Psychology, University of Sydney, NSW 2006, Australia.

Commercial relationships: none.

one would expect from probability distributions with equalized, or “white,” spatio-temporal spectra. This naturally occurring, low spatial and temporal frequency-biased relationship with energy, known as an inverse amplitude spectrum, has a characteristic slope (linear in log–log coordinates) that is more or less invariant across environments, scales, and studies ( $f^{\text{spatial}}, f^{\text{temporal}}, \approx -1$ ) (Bex, Dakin, & Mareschal, 2005; Burton & Moorhead, 1987; Dong & Atick, 1995; Field, 1987; Field & Brady, 1997; van Hateren & van der Schaaf, 1996; but see Tolhurst, Tadmor, & Chao, 1992). Examples of such low-frequency-biased luminance distributions are illustrated spatially in Figure 1b (pink and red-brown insets) and temporally in Movie 1. How does the visual system respond to this naturally occurring low-frequency bias?

## Experiment 1: Point of subjective equalization—spatial frequency

In the spatial domain, psychophysical demonstration indicates that the visual system effectively equalizes its perceptual response to low spatial frequency-biased ( $SF^{-1}$ ) input, a phenomenon known as whitening (Brady & Field, 1995, 2000; Field, 1987; Field & Brady, 1997). This can be observed phenomenologically in Figure 1b, which depicts three sets of visual noise distinguished by the spatial distribution of their amplitude spectra. Despite containing an approximately equal proportion of spatial frequencies (within Nyquist limits), “white noise” patterns ( $\alpha = 0$ ) appear predominantly high frequency biased (see white inset pattern in Figure 1b). These white noise spatial patterns become more balanced perceptually by imposing a negative exponential bias ( $\alpha < 0$ ) to their amplitude spectra (a form of low-pass filtering; see pink and red-brown insets in Figure 1b; Field & Brady, 1997). Remarkably, perceptually balanced percepts tend to occur when spectral exponents approximate those observed in natural scenes ( $\alpha \approx -1$ ). We quantified this effect psychophysically by presenting subjects with a series of static achromatic noise patterns of variable spatial exponent (see Methods) and asked them to indicate any bias they observed in spatial frequency content (i.e., low vs. high frequency dominant).

### Methods

**Materials**—In all experiments, stimulus presentation was driven by an ATI Radeon X1600 graphics card, generated using the Psychophysics Toolbox version 3 software (Brainard, 1997) and displayed on a Mitsubishi Diamond Pro Monitor (1024 × 768 pixel resolution in a display area of 24 × 24 cm, 100 Hz vertical refresh, mean luminance = 52 cd/m<sup>2</sup>) with a linearized gamma. All experiments employed 10.8-bit luminance resolution achieved using bit stealing. Viewing was binocular and fixed at a viewing distance of 57 cm using a chin rest.

**Stimuli**—The stimuli in Experiment 1 were spatially broadband achromatic noise patterns presented within a circular aperture (diameter = 6 degrees of visual angle) centered on fixation. Aperture edges were smoothed with a raised cosine ramp (standard deviation = 20 pixels) against a gray background held at mean luminance. Noise patterns were independently created on each separate trial by assigning each pixel within a 256 × 256 matrix, a luminance value derived from a uniformly random distribution of values between  $-1$  and  $1$ . The DC component of this distribution was set to zero and later rescaled to mean luminance (52 cd/m<sup>2</sup>) to ensure that the mean luminance of each image was identical. The fast Fourier transform (FFT) was calculated for each noise image and was averaged across all orientations. The amplitude spectrum ( $A_f$ ) of each image decayed as a function of spatial frequency (SF) with exponent ( $\alpha$ ):

$$A_{\text{SF}} = 1/\text{SF}^{\alpha}. \quad (1)$$

The spatial frequency exponent ( $\alpha$ ) was varied across trials and was the dependent variable (see Figure 1). Each stimulus pattern had an RMS contrast of 30% calculated after the FFT filtering operation. The total duration of each stimulus pattern was 1 second and was ramped on and off using a raised cosine ( $SD = 50$  ms).

**Procedure**—An interleaved adaptive staircase procedure was used to estimate the spatial exponent at which observers perceived an equal proportion of “low” and “high” spatial frequency energy. The point of subjective equalization was determined by presenting subjects with test stimuli of variable exponent ( $\alpha$ ) across trials and asking them to report, using a keypress, whether the stimulus was perceived to possess either a higher proportion of “low” or “high” spatial frequency energy. The physical exponent of each stimulus was varied in response to the subjects’ report using two randomly interleaved adaptive staircases (starting points:  $\alpha = .2$  and  $-1.8 \pm .2$ ; step size = 0.1). Each staircase contained 50 trials, and each naïve participant repeated the experiment once (a total of 200 trials). Subject JC repeated the experiment 11 times (1200 trials). Prior to the experiment, naïve subjects were each presented with a printed array of spatially narrowband (1 octave) filtered noise patterns, each centered on a different spatial frequency ranging from 0.5 to 16 cycles/image in 1 octave steps. This procedure was employed to familiarize each subject visually with the concept of spatial frequency—the dimension relevant to the task—and to ensure that they could reliably differentiate and identify low and high spatial frequency images. By fitting a psychometric function to each observer’s data, we calculated the point of subjective equalization to be the exponent ( $\alpha$ ), which generates an equal proportion (50%) of “low” and “high” “spatial frequency dominant” responses.

**Observers**—Four experienced psychophysical observers with normal vision participated in the experiment. Three were naïve to the purposes of the study. Author JC (black data and curve in Figure 2a) also served as a subject and was aware of the study’s hypotheses.

## Results and discussion

As can be seen in Figure 2, flatter (i.e., whiter) spectral distributions with slopes closer to zero resulted in proportionally fewer “low spatial frequency dominant” responses. This was invariably so at  $\alpha$  values between 0 and  $-.6$ . Conversely, observers invariably reported  $\alpha$  values less than  $-1.2$  to be low spatial frequency dominant. Critically, perceptual equalization occurred for exponents ranging between  $-.88$  and  $-.93$ . This range is significantly different from physical equalization ( $p < .001$ ) and well within the range measured in natural images (Brady & Field, 1995; Tolhurst et al., 1992).

It has been argued that spatial whitening confers both computational and metabolic advantage by equalizing its response to the low-frequency bias present in natural images, thereby reducing input redundancy and increasing informational efficiency (Atick & Redlich, 1992; Barlow, 2001; Simoncelli & Olshausen, 2001). Given that natural image sequences exhibit  $f^{-1}$  amplitude spectra in both spatial frequency and temporal frequency dimensions, this begs the question of whether there exists a temporal whitening/equalization process analogous to that observed in the spatial frequency dimension.

## Experiment 2: Point of subjective equalization—temporal frequency

We addressed the question using an analogous procedure to that used in Experiment 1. We asked subjects to identify whether they perceived either a greater proportion of low temporal frequency or high temporal frequency energy within temporally broadband noise sequences of variable temporal exponent ( $\beta$ ). By systematically varying  $\beta$  in each stimulus using an adaptive staircase procedure (see Methods; shallower negative exponents possess a greater proportion of high temporal frequencies) and fitting a psychometric function to observers’

responses (Figure 3a), we estimated the distribution of energy required to generate a percept that was spectrally unbiased temporally (i.e., equalized).

## Methods

**Stimuli**—The stimuli used in Experiment 2 were spatially narrowband and temporally broadband achromatic noise patterns presented within a circular aperture (diameter = 6 degrees of visual angle) with a cosine-ramped outer edge (standard deviation = 20 pixels) against a gray background held at mean luminance.

Noise sequences were created independently for each trial by assigning each pixel within a  $128 \times 128 \times 64$  matrix, a luminance value derived from a uniformly random distribution of values between  $-1$  and  $1$ . The DC component of this distribution was set to zero and later rescaled to mean luminance ( $52 \text{ cd/m}^2$ ) to ensure that the mean luminance of each image sequence was identical. The fast Fourier transform (FFT) was calculated for each image within each movie sequence image (i.e., spatially) and between images within each sequence (i.e., temporally). The spatial amplitude spectrum ( $A_{SF}$ ) of each image was band-pass ( $SD = 1/2$  octave) centered at  $1.6$  cycles/deg. The temporal amplitude spectrum decayed exponentially with exponent ( $\beta$ ):

$$A_{TF} = 1/TF^\beta. \quad (2)$$

The temporal frequency exponent ( $\beta$ ) was varied across trials and was the dependent variable in this experiment (see Movie 1). Each stimulus sequence had an RMS contrast of 30%. The total duration of each stimulus pattern was 1 second and was ramped on and off using a raised cosine ( $SD = 50$  ms). Because of the computational demands of our 3-D Fourier filtering, in this experiment we reduced the size of the  $x, y$  pixel matrix used in Experiment 1 by a factor of two. To preserve stimulus visual angle across experiments, we increased the spatial extent (while decreasing spatial resolution) of the filtered stimulus by a factor of two using linear interpolation between pixels.

**Procedure**—An adaptive staircase procedure was used to estimate the test stimulus amplitude exponent in which observers perceived an equal proportion of low and high spatial frequency energy. The point of subjective equalization was determined presenting subjects with a series of stimuli in which the temporal exponent ( $\beta$ ) varied from trial to trial and asking them to report whether trial was perceived to be either “low” or “high” temporal frequency dominant. The physical exponent of each stimulus was varied in response to the subjects’ report using two randomly interleaved adaptive staircases (starting points:  $\beta = .0$  and  $-2 \pm .2$ ; step size =  $0.1$ ). Each staircase contained 50 trials, and all but one participant repeated the experiment once (a total of 200 trials). Author JC repeated the experiment seven times (black curve, Figure 3a). Prior to the experiment, naïve participants were presented with two spatially and temporally narrowband (0.5 octaves above and below center frequency) filtered noise patterns centered on 1 and 10 Hz both at 2 cycles/deg (diameter = 6 degrees of visual angle). This procedure was employed to familiarize each subject visually with the concept of temporal frequency—the dimension relevant to the task—and to ensure that they could reliably differentiate and identify low and high temporal frequency images. An initial run of 20 trials was presented to each naïve participant whereby 1- and 10-Hz reference stimuli were presented simultaneously with the test stimulus, centered 10.4 degrees to the left (1 Hz) and right (10 Hz) of fixation. By fitting a psychometric function to each observer’s data, we calculated the point of subjective equalization to be that exponent ( $\beta$ ) that generates an equal (i.e., unbiased) proportion of “low” and “high” “temporal frequency dominant” responses.

**Observers**—Nine psychophysical observers with normal vision participated in the experiment. Seven were naïve to the Experiment’s purposes and two were familiar with the hypotheses.

## Results and discussion

As can be seen in Figures 3a–3c, observers uniformly reported that stimuli composed of unbiased temporal spectra ( $\beta \approx 0$ ) contained a greater proportion of “fast” (high temporal frequency) compared with “slow” (low temporal frequency) luminance modulation (see Movie 1a). Moreover, exponents ranging between  $\approx -0.7$  and  $-1.5$  (mean =  $-0.97$ ) were reported to contain an equal (i.e., unbiased) proportion of low and high temporal frequency modulation (Figures 3a, 3b, and 3d; see Movie 1b). This latter range of exponents corresponds to the range of temporal amplitude spectra measured in natural scenes (Dong & Atick, 1995; van Hateren & van der Schaaf, 1996). This experiment provides the first direct demonstration that the human visual system not only whitens its temporal input but it does so via a transient response bias of a magnitude that appears to compensate for the low temporal frequency bias present in natural scenes.

## Experiment 3: Effect of spatial frequency on point of subjective temporal equalization

Thus far, we have investigated how the visual system responds to broadband temporal frequency within a narrow (1 octave) range of spatial frequencies centered around 1.6 cycles/deg. In natural viewing conditions, small fixational instability generate proportionally more temporal luminance variation at fine spatial scales (high spatial frequencies) than at more coarse scales (lower spatial frequencies; Rucci, Iovin, Poletti, & Santini, 2007). We conducted a third experiment to examine how the visual system responds to this interaction between stimulus spatial and temporal frequency dimensions. By measuring the temporal exponents at which subjects perceive an equalized proportion of broadband temporal frequency energy across a broad range of narrowband spatial frequencies (0.5, 1, 2, 4, and 8 cycles/deg), we seek to determine whether the visual system’s transient response bias is affected by the absence of high spatial frequency structure in the test stimulus used in Experiment 2. If the subjectively equalized temporal exponents measured in Experiment 2 do indeed reflect a whitening process optimized to the statistical properties of natural viewing (including fixational eye movements; Rucci et al., 2007), the increase in eye-movement-induced high temporal frequencies at finer spatial scales under natural viewing conditions predicts that the transient response bias observed in Experiment 2 (inferred from the subjectively equalized exponent,  $\beta \approx -1$ ) will decrease as spatial frequencies increase. That is to say, the exponent of subjective equalization ( $\beta$ ) is expected to increase (i.e., be closer to zero) at higher spatial frequencies.

## Methods

We conducted an experiment using an identical temporal equalization task to that used in Experiment 2. In this experiment, however, we varied the center spatial frequency around which the stimuli were band-pass filtered ( $SD = 1/2$  octave, centered around 0.25, 0.5, 1, 2, 4, and 8 cycles/deg). Three subjects participated, two authors (JC and DA) and one who was naïve to the purposes of the experiment and did not participate in Experiment 2. Each subject measured their point of subjective temporal equalization for each spatial frequency during separate blocks using two adaptive randomly interleaved staircases, each consisting of 60 trials. Block order was randomized for each subject. A LaCie electron 22 blue III 100 Hz monitor with a mean luminance of 28.4 cd/m<sup>2</sup> was used to display stimuli in this experiment.

## Results and discussion

As can be seen in Figure 4, our estimates reveal that all subjects chose exponents  $\beta < 0$  to be neither too fast, nor too slow, and therefore equalized in their perceived temporal frequency distribution. This response bias indicates that subjects either underestimate the proportion of low temporal frequencies in the display and/or overestimate the more transient stimulus energy. Critically, Figure 4 shows that the exponent of subjective temporal equalization increased as a function of stimulus spatial frequency. This result suggests that the visual system becomes progressively less transiently biased in its response at higher spatial frequencies. It is worth noting that natural viewing conditions produce amplitude spectra with a similar, although inverted, spatio-temporal coupling to those derived perceptually in Experiment 3. Small transient fixational eye movements produce small displacements in the retinal image. Consequently, fine spatial scales (high spatial frequencies) carry proportionally more high temporal frequency energy than lower spatial frequencies (Rucci et al., 2007). If the objective of the transient bias observed in Experiments 2 and 3 is to whiten the visual system's response to the low temporal frequency bias observed in natural scenes, the greater prevalence of high temporal frequency energy at high spatial frequencies (Rucci et al., 2007) predicts that the transient response bias should decrease with increasing spatial frequency. This is what we observe.

In addition to the reduction in transient response bias observed at higher spatial frequencies, a similar reduction appears to be evident at the lowest spatial frequency tested (mean  $\beta_{\text{perceived}} \approx -.79$  at .25 cycles/deg). Why this reduction in transient bias should occur is unknown as it does not appear to correlate with any increase in naturally occurring high temporal frequency energy at this relatively coarse spatial scale (which actually appears more low frequency biased ( $1/\text{TF}^2$ ); Dong & Atick, 1995). One possibility is that this spatial tuning profile may simply reflect the greater proportion of dynamically driven V1 neurons optimized to intermediate spatial frequencies. Alternatively, a more complex interaction may account for this effect such as low spatial frequency suppression combined with higher frequency excitation reported in macaque V1 (Bredfeldt & Ringach, 2002). Future research is required to determine the validity of these hypotheses.

It is tempting to speculate that the temporal whitening effect observed in Experiments 2 and 3 may result from a similarly early transient bias. Both retinal ganglion and lateral geniculate neurons in the magnocellular pathways of cats and primates are known to exhibit a similar transient/low spatial frequency response (Derrington & Fuchs, 1978, 1979; Enroth-Cugell & Robson, 1966; Lee, Pokorny, Smith, Martin, & Valberg, 1990). Although psychophysical estimates of sensitivity ( $1/\text{detection threshold}$ ) to various rates of temporal modulation (known as the modulation transfer or deLange function) usually possess only a moderate transient bias (Bex & Langley, 2007; Hess & Snowden, 1992; see Figure 5a), supra-threshold matching studies indicate a progressively more band-pass high contrast, high luminance-driven transient response (Bex & Langley, 2007; Georgeson, 1990; Georgeson & Harris, 1990). It is possible, therefore, that the transient bias observed in Experiments 2 and 3 may simply result from the supra-threshold response of a single transient filter within each spatial channel.

Other evidence provides a more complex picture. The standard model assumes that the modulation transfer function represents the combined output of multiple, independent temporal frequency-selective mechanisms or channels: one low-pass (or "sustained") and one (or two) more transient and band-pass (Anderson & Burr, 1985, 1989; Boynton & Foley, 1999; Cass & Alais, 2006; Foley & Boynton, 1993; Hess & Snowden, 1992; Hess, Waugh, & Nordby, 1996; Lehky, 1985; Meese & Holmes, 2007; Meier & Carandini, 2002). These hypothetical channels are evident in the fitted modulation transfer function in Figure 5a (black curve), which represents linear summation of a pair of Gaussian functions (i.e., the

purported channels, see gray curves). In principle, the idea that multiple, independent, and bandwidth-limited channels underlie sensory processing along the temporal frequency dimension is analogous to the dominant view of spatial frequency processing in the primate visual system (Anderson & Burr, 1985; Blakemore & Campbell, 1969a, 1969b; Campbell & Robson, 1968). According to this multiple channels account, whitening could result either from interactions between channels and/or differential channel gains and/or saturation constants.

An alternative explanation, originally proposed to account for spatial whitening relates to channel bandwidth. Because the bandwidths of spatial frequency channels increase as a function of (linearly scaled) frequency, the greater integration afforded by the higher frequency filters predicts a proportional increase in the population response in favor of higher frequencies (Brady & Field, 1995). Applying this principle to the exponential decay observed in natural temporal amplitude spectra predicts a whitening (flattening) of the resulting integrated response. Although, in principle, both the integration and the differential gain/saturation accounts of temporal whitening may predict different outcomes with respect to stimulus contrast, both assume that sustained and transient channels operate independently. For the purposes of this paper, we refer to both the integration and the differential gain/saturation accounts collectively as the *independent channels account of temporal whitening*.

#### Experiment 4: Temporal frequency masking

Early psychophysical masking studies revealed the existence of two independent temporal channels, one low-pass and “sustained” and one or two more transient band-pass channels (Anderson & Burr, 1985; Hess & Snowden, 1992; Lehky, 1985; Snowden & Hess, 1992). Temporal masking maps the extent to which observers’ sensitivity to a particular (target) temporal frequency is attenuated by the simultaneous and spatial superimposition of additional, target-irrelevant (masking) frequencies. Masking stimuli whose spectral energy falls within the bandwidth of the target-sensitive channel generate a target-irrelevant increase in the channel’s response. This is equivalent to reducing the target-sensitive channel’s signal-to-noise ratio, thereby reducing target detectability. Conversely, masking stimuli whose spectral energy falls beyond the bandwidth of the target-sensitive channel contribute no appreciable increase in the proportion of noise associated with its total response, exerting minimal reduction in target sensitivity. Systematic examination of the rise and fall of target sensitivity incurred as a function of masking frequency can be used to estimate the number, locations, and shapes of the underlying channels (Anderson & Burr, 1985; Cass & Alais, 2006; Hess & Snowden, 1992; Lehky, 1985; Snowden & Hess, 1992).

By using a higher resolution masking procedure to that employed previously, Cass and Alais (2006) this strong registration, Cass and Alais concluded that the channels were not independent and that the low-frequency channel receives masking input from the high-frequency channel, but not vice versa.

The phenomenal implications of this asymmetry in temporal masking are illustrated in Movie 2. Whereas both low- and high-frequency masks readily disrupt perception of low-frequency target modulations, fast modulations are relatively immune to low-frequency masks. This perceptual asymmetry in favor of transient modulation suggests that masking and temporal whitening may recruit similar, possibly common, mechanisms. If the high-frequency perceptual dominance observed in the whitening phenomenon (Experiment 2) is in fact due to transient-biased channel interactions observed in masking (Cass & Alais, 2006), then the magnitude of any masking-induced transient bias should be of an order that predicts whitening.

## Methods

**Stimuli**—The stimuli used in Experiment 4 were spatio-temporally narrowband achromatic noise sequences presented within a circular aperture (diameter = 6 degrees of visual angle) with a cosine-ramped outer edge (standard deviation = 20 pixels) against a gray background held at mean luminance. Target and masking stimuli were each created from independent noise sources on each trial. Noise was generated by assigning each pixel within a  $128 \times 128 \times 64$  matrix, a luminance value derived from a uniformly random distribution of values between  $-1$  and  $1$ . The DC component of this distribution was set to zero and later rescaled to mean luminance ( $52 \text{ cd/m}^2$ ) to ensure that the mean luminance of each image sequence was identical. The fast Fourier transform (FFT) was calculated for each image within each movie sequence image (i.e., spatially) and between images within each sequence (i.e., temporally). The spatio-temporal amplitude spectrum of each image was band-pass ( $SD = 1/2$  octave) centered at  $2.4$  cycles/deg. The center temporal frequency of target and masking stimuli were varied independently of each other, with particular combinations blocked (in pseudo-random order) across trials. Target temporal frequencies were centered on either  $1$  or  $15$  Hz, with masking stimuli centered in approximately half octave steps between  $1$  and  $60$  Hz. The total duration of each stimulus pattern was  $1$  second and was ramped on and off using a raised cosine ( $SD = 50$  ms).

**Procedure**—The experimental methodology, based upon Cass and Alais (2006) and Hess and Snowden (1992), was conducted in two stages. The first stage involved measuring target detection thresholds in the absence of masking stimuli. In the second stage, target contrasts were fixed at  $4$  dB above threshold then superimposed (added) with one of two identical masking stimuli whose contrast varied across trials. The temporal frequencies of target and masking stimuli were chosen pseudorandomly and were blocked across trials. A spatial two-alternative forced-choice procedure was used to estimate threshold. On any given trial, the target stimulus appeared either  $5.2$  degrees above or below fixation and subjects were required to report which of these two locations the target had been presented. Target location was randomized across trials. In masking conditions, identical masking stimuli were presented simultaneously at each potential target location. In each block of trials, two randomly interleaved adaptive staircases procedures were used to estimate the detection threshold ( $60$  trials/staircase). This was repeated once at each mask and/or target frequency combination. In target-only trials, correct and incorrect responses elicited subsequent decreases and increases in target contrast respectively. In target + mask trials, correct responses elicited subsequent increases in mask contrast, with incorrect trials producing decreases in mask contrast. Corrective feedback was provided following each trial by briefly changing the color of the fixation point (red = incorrect; blue = correct).

**Observers**—Two psychophysical observers with normal vision participated in the experiment. Both were experienced psychophysical observers. One was naïve to the experiment's purposes and the other (author JC) was aware of the experiment's hypotheses.

## Results and discussion

As can be observed directly in Movie 2 and indirectly in the masking data in Figure 6, there is a compelling asymmetry in the perceptual effects of low versus high temporal frequency masking on sensitivity to  $1$ - versus  $15$ -Hz targets. We find that while lower rates of temporal modulation ( $<4$  Hz) tend to interfere with the detectability of  $1$ -Hz targets, with little or no effect on  $15$ -Hz targets,  $1$ - and  $15$ -Hz targets alike are robustly masked by modulation rates  $>4$  Hz, peaking at around  $19$  Hz. While this replicates the basic qualitative results of our earlier masking study, some differences do emerge. Most notable is a rightward shift in center frequency of both the more transient peak of the bimodal Gaussian (derived using  $1$ -Hz targets) and the unimodal transient band-pass function (derived using  $15$ -Hz targets)



from 8 to 12 Hz (Cass & Alais, 2006) to  $\approx 20$  Hz. A potential explanation for this transient shift may be that the current experiment used parafoveal presentation of test and masking stimuli rather than the foveal stimulation employed in our original study (Snowden & Hess, 1992). Despite this difference, both studies share the same qualitative features. Whereas the masking function derived using 15-Hz targets is well fitted by a single high band-pass Gaussian (peaking at  $\approx 20$  Hz), 1-Hz masking is better characterized by the sum of two band-pass Gaussians (peaking at  $\approx 1$  and 20 Hz; red curve). Cass and Alais (2006) interpreted the spectral registration between the high band-pass and higher secondary peaks (respectively derived using 1- and 15-Hz targets) as evidence that the underlying temporal frequency channels interact asymmetrically, such that the lower temporal frequency channel receives masking input from the higher temporal frequency channel.

By assuming that this asymmetric interaction between temporal channels is inhibitory in nature (Allison, Smith, & Bonds, 2001; Morrone, Burr, & Speed, 1987) (see Figure 6b), Cass and Alais (2006) showed that summing the inferred inhibitory and excitatory outputs of these channels produced a high temporal frequency-biased response (see Figure 6c). The response of this system to broadband noise can be estimated via convolution. Convoluting this high temporal frequency response with an inverse amplitude spectrum ( $\beta = -1$ ) strongly flattens the resultant curve between 1 and 15 Hz (see Figure 6d). We propose that this flattening represents a functional transformation between the  $1/f$  input associated with natural scenes and the visual system's perceptual output. The strong relationship between the predicted perceptual response of the *asymmetric inhibition model* to broadband noise (Cass & Alais, 2006) and the psychophysical results in Experiment 2 strongly suggests it as a candidate mechanism for temporal whitening; a process that flattens the spectrum so all frequencies are effectively present in the resultant percept. This example of perceptual whitening in the temporal domain may be distinguished from earlier accounts of whitening in the spatial domain (Brady & Field, 1995) in which spatial frequency channel responses themselves are equalized in response to  $1/f$  spatial input. Because the visual system's output in the spatial domain is simply the sum of independent channels, equalizing the channel outputs is the simplest way to achieve whitening. This does not apply to the temporal domain if temporal channels are linked via an asymmetric inhibition. The important point, however, is that both accounts are equivalent functionally, in that the response to natural (i.e.,  $1/f$ ) input is flattened.

There are, however, other mechanisms that may flatten the  $1/f$  temporal amplitude spectrum. According to channel theory, masking is caused by a reduction in the signal-to-noise ratio associated with a signal-relevant channel's response. In masking, signal-to-noise ratios may be reduced by either reducing the signal (via an active inhibitory process; Cass & Alais, 2006; Graham, Chandler, & Field, 2006; Srinivasan, Laughlin, & Dubs, 1982; or an inactive one, e.g., *synaptic depression*; Chance, Nelson, & Abbott, 1998) and/or increasing the noise. Temporal whitening could therefore, in principle, be due to the low temporal frequency channel being more susceptible to high temporal frequency-driven noise than the converse situation (see Figure 6b). One can conceive of several processing architectures capable of producing such an asymmetric noisy interaction. The output of high temporal frequency-biased mechanisms may, for example, disrupt the efficiency that low temporal frequency information is encoded and/or transmitted. This may be accomplished by noisy interactions between otherwise parallel pathways (e.g., transient and sustained channels) or else temporal mechanisms with differential susceptibilities to input noise from a common source (Chance et al., 1998; Langley & Bex, 2007; Wang, Liu, Sanchez-Vives, & McCormick, 2003). We refer to this general class of models as the *asymmetric noise hypothesis*.

$$\text{ASoG}_{\text{Masking}} = \left( a_1 \cdot \exp \left[ -\frac{(x - .1_{\text{Hz}})^2}{2 \cdot b_1^2} \right] \right) - \left( a_2 \cdot \exp \left[ -\frac{(x - .1_{\text{Hz}})^2}{2 \cdot b_2^2} \right] \right) + \left( a_3 \cdot \exp \left[ -\frac{(x - 19_{\text{Hz}})^2}{2 \cdot b_3^2} \right] \right). \quad (3)$$

Equation 3: asymmetric inhibition and asymmetric noise model of masking, where  $a_{1\&2}$  = amplitude and  $b_{1\&2}$  = bandwidth of Gaussian peaks (low and high temporal frequency) derived from masking of 1-Hz target (red curve in Figure 6a);  $a_3$  = amplitude and  $b_3$  = bandwidth of Gaussian peak derived from masking of 15-Hz target (blue curve in Figure 6a). Note that the subtractive term between first and second Gaussian components results from an active inhibitory process in the asymmetric inhibition model and from input noise in the asymmetric noise model. See Figure 6b for graphical representation.

## Experiment 5: Temporal frequency adaptation

To empirically compare these various accounts of temporal whitening, *independent channels*, *asymmetric inhibition*, and *asymmetric noise*, we conducted a second experiment using a psychophysical adaptation paradigm. Adaptation refers to the perceptual and/or neurophysiological effects of prior stimulation on the response to subsequently presented test stimuli. Adaptation effects are generally thought to be the result of attenuative gain control processes that serve to renormalize the sensory system's response to prolonged biased input (Clifford et al., 2007; Kohn, 2007), thereby reducing input redundancy. According to this approach, changes in sensitivity are generally assumed to be proportional to the responsiveness of the channel(s) to the adapting stimulus. Whereas prolonged stimulation of an excitatory mechanism is believed to produce a subsequent reduction in its response (suppression), adapting an inhibitory mechanism has been found to reduce its inhibitory effect (disinhibition) (Durand, Freeman, & Carandini, 2007; Kohn & Movshon, 2003). Based on these principles, each of the temporal whitening models described above is expected to generate the following distinct patterns of adaptation performance:

The *independent channels* model operates via independent excitatory processes with distinct temporal frequency selectivities and bandwidths. According to this view, adapting to a high range of temporal frequencies (>6 Hz) is predicted to elevate thresholds within this range, with less effect on thresholds to lower temporal frequencies. Conversely, low temporal frequency adaptation (<6 Hz) is predicted to elevate low temporal frequency thresholds with less effect on thresholds for more transient stimuli (see Figures 7a–7c).

The *asymmetric inhibition* model depends upon unilateral high-frequency-driven inhibition of the low temporal frequency channel (see Figure 7d). According to the principle of adaptive normalization, prolonged exposure to high temporal frequencies should induce subsequent (i) high temporal frequency suppression (threshold elevation; Figure 7e) and (ii) disinhibition of the low temporal frequency channel (reduced threshold elevation, possibly even threshold reduction, i.e., facilitation; Figure 7f). Conversely, low temporal frequency adaptation is predicted to elevate low temporal frequency thresholds (suppression) while exerting little or no suppressive effect on high temporal frequency thresholds.

The *asymmetric noise* account of temporal whitening operates via noisy interaction feeding from high-frequency-biased to low-frequency-biased perceptual mechanisms, but not vice versa (see Figure 7g). Consequently, high temporal frequency adaptation is predicted to induce subsequent increases in low and high temporal frequency detection thresholds (suppression) (Figure 7h). In contrast, low temporal frequency adaptation is predicted to increase low temporal frequency thresholds, with less effect on high-frequency thresholds (Figure 7i).

## Methods

**Stimuli**—The target and adapting stimuli in used in Experiment 5 were spatio-temporally filtered noise sequences presented within a circular aperture (target diameter = 6 degrees of visual angle; adapting diameter = 7 degrees of visual angle) with a cosine-ramped outer edge (standard deviation = 20 pixels) against a gray background held at mean luminance. Noise generation and spatio-temporal filtering procedures were identical to those used in Experiment 4. Spatial frequency was band-pass ( $SD = 1/2$  octave) centered at 2.6 cycles/deg. Target temporal frequencies were centered on either 1 or 15 Hz. Adapting stimuli were always presented at 30% RMS contrast and were centered at 1, 2, 4, 6, 8, 10, 15, 20, 40, and 60 Hz. The total duration of each stimulus pattern was 1 second. Adapting stimuli were presented for 60 seconds on the first trial in a given block, then 10 seconds on each subsequent trial. Target and masking stimuli were ramped on and off using a raised cosine ( $SD = 50$  ms).

**Procedure**—The experimental methodology consisted of two stages. The first stage involved measuring target detection thresholds prior to adaptation. Target detection thresholds were measured using an identical spatial 2AFC procedure to that described in Experiment 4, in which target stimuli would appear 5.6 degrees above or below fixation. A staircase controlled target contrast, increasing contrast in response to localization errors and decreasing following correct localization for a total of 40 trials. All threshold estimates were based on two separate staircases. Each block of 40 trials (i.e., a single staircase) involved a single target and adapting temporal frequency. The temporal frequencies of adapting and/or test stimuli used in a given block were chosen in pseudorandom order. Adapting stimuli were simultaneously presented 5.6 degrees above and below fixation. Target onsets initiated on the frame immediately following the termination of the adapting stimulus. On any given trial the target stimulus appeared either 5.2 degrees above or below fixation and subjects were required to report which of these two locations the target had been presented. Target location was randomized across trials. Corrective feedback was provided following each trial by briefly changing the color of the fixation point (red = incorrect; blue = correct).

**Observers**—Two psychophysical observers with normal vision participated in the experiment. Both were experienced psychophysical observers. One was naïve to the experiment's purposes and the other (author JC) was aware of the experimental hypotheses.

## Results and discussion

The results of our temporal frequency adaptation experiment (Figure 8a) are clear. Whereas sensitivity to low temporal frequency targets (1 and 4 Hz) is robustly reduced following adaptation to similarly low temporal frequencies (1–4 Hz), sensitivity to higher temporal frequency targets (10 and 30 Hz) is relatively less affected. This is to be distinguished from the strong sensitivity reducing effects of >6 Hz adaptation on all targets regardless of their temporal frequency. Both of these results are reminiscent of an early study by Pantle (1971) and, more recently, Langley and Bex (2007). This asymmetry in the effects of low and high temporal frequency adaptation on the subsequent sensitivity to low and high temporal frequency targets is illustrated by both the shapes and the locations of the fitted Gaussian peaks (Figure 8a). Whereas adaptation of low temporal frequency targets (1 Hz) is well fitted by a multimodal Gaussian distribution with peaks distributed across the temporal spectrum (approximately 1 and 20 Hz (JC) and 0.1 and 10 Hz (AP); see red curves), the adaptation peaks derived using 15-Hz targets exhibit an overall transient shift and are well fitted by a unimodal Gaussian centered on approximately 20 Hz.

These results are not predicted by either *independent channels* or *asymmetric inhibition* accounts of temporal whitening but do accord well with the *asymmetric noise* account.

Because high (but not low) temporal frequency adaptation induces threshold elevations for both low and high temporal frequencies alike, this implies that lower frequency-selective channel receives a greater proportion of (noisy) adaptation-inducing high temporal frequency input than the converse situation. This is modeled in Figure 8b, with parameters derived from adaptation functions based on sensitivity to 1-Hz targets (red curves in Figure 8a). Assuming that sensitivity to a particular target frequency is a function of the output of a Gaussian-shaped channel whose peak is most proximal to this frequency (1 Hz for JC and 0.1 Hz for AP), we center the bandwidths and amplitudes of all other Gaussians derived from this adaptation function. An equivalent procedure is repeated for the adaptation function derived using 15-Hz targets (blue curve, Figure 8b), which for these two subjects involves a single Gaussian peak centered at 19 Hz. These peak registration processes are then followed by a linear summation of all Gaussians anchored to the two spectral locations. This procedure is depicted symbolically in Equations 3 and 4.

$$\text{ASoG}_{\text{Adaptation}} = \left( a_1 \cdot \exp \left[ -\frac{(x - 0.1_{\text{Hz}})^2}{2 \cdot b_1^2} \right] \right) + \left( a_2 \cdot \exp \left[ -\frac{(x - 0.1_{\text{Hz}})^2}{2 \cdot b_2^2} \right] \right) + \left( a_3 \cdot \exp \left[ -\frac{(x - 19_{\text{Hz}})^2}{2 \cdot b_3^2} \right] \right). \quad (4)$$

Equation 4: asymmetric noise model of adaptation; where  $a_{1\&2}$  = amplitude and  $b_{1\&2}$  = bandwidth of Gaussian peaks (lowest to highest temporal frequency) derived from adaptation function for 1-Hz target (red curve in Figure 8a) and  $a_3$  = amplitude and  $b_3$  = bandwidth of Gaussian peak derived from adaptation function for 15-Hz target (blue curve in Figure 8a). See Figures 8b and 8c for graphical representation.

As can be seen in Figure 8c, recombining the adaptation-derived temporal channels in this asymmetric fashion results in a transient bias. Convolving this transient bias with an inverse temporal amplitude spectrum ( $\beta = -1$ ) generates a general flattening in its response (see red-blue dashed curve in Figure 8d). By comparing the magnitude of this flattening effect with that produced by the un adapted modulation transfer function (black curve in Figure 8d), we observe that this effect is substantial within a different spectral range for each subject (2–10 Hz for JC; 1–2 Hz for AP). Despite these differences, the fact that adaptation appears to induce an overall more transient response bias suggests that adaptation itself may represent a form of temporal whitening in the human visual system. If one of the goals of sensory encoding and transmission is in fact to reduce input redundancies, then the observed reduction in sensitivity to the naturally more prevalent low temporal frequencies seems consistent with this objective.

**Relationship to the Motion Aftereffect (MAE)**—Curiously, the asymmetry in temporal frequency-contingent contrast adaptation we observe is not apparent in studies that measure changes in perceived direction of translating motion which result from prolonged adaptation to translating luminance patterns (known as the motion aftereffect (MAE); Alais, Verstraten, & Burr, 2005; van der Smagt, Verstraten, & van de Grind, 1999). In these MAE studies, adapting to high drift rates (>~20 Hz) elicits changes in the perceived direction of subsequently presented dynamic test patterns, with significantly less direction aftereffect observed when using static or low temporal frequency test patterns. Conversely, low drift rates are found to induce significantly greater MAEs when employing static or low temporal frequency test patterns compared with highly dynamic tests. The rate specificity of the static and dynamic MAEs implies that these effects result from differential yet independent adaptation of low and high temporal frequency (or speed)-selective channels.

Why the MAE should fail to generate an analogous asymmetry to the temporal frequency-contingent contrast adaptation we observe is unknown. One possibility is that the MAE and

contrast adaptation may result from distinct mechanisms, with the MAE possibly occurring at a higher level of processing. An alternative possibility relates to the broadband spatio-temporal structure of the test and/or adapting patterns used in the MAE studies. That the spatio-temporal frequency structure of our stimuli was precisely defined as narrowband undermines direct comparison of the results from these two paradigms. Future experimentation is required using spatio-temporally equivalent stimuli to determine the nature of any relationship between the temporal frequency determinants of contrast adaptation and the MAE.

**Relationship between temporal masking and adaptation**—Comparing the results of our masking and adaptation experiments reveals both divergent and complementary effects. Qualitatively, both masking and adaptation yield two similar sets of functions: a multimodal function, derived using 1-Hz targets, and a transient band-pass function derived from high-frequency targets, centered at approximately 20 Hz. The strong correspondence between the transient band-pass functions derived using these two methods suggests a common mechanism. What is perhaps most problematic for the idea that masking and adaptation are mediated by equivalent mechanisms, however, is the lack of spectral registration between the transient bandpass peak (blue curves) and the transient secondary peaks (red curves) observed in the adaptation functions derived from subject AP (see Figure 8a). This is distinguishable from masking, which exhibits a strong registration under analogous conditions. One should therefore consider the possibility that transient adaptation of 1-Hz targets is dissociable from the mechanisms mediating transient adaptation of 15-Hz targets. This, of course, would imply that our asymmetric model does not generalize to adaptation. An alternative explanation (that is consistent with our asymmetric noise model), however, might be that the apparently unimodal, band-pass transient adaptation function derived using high temporal frequency targets (see blue curve, Figure 8a) may in fact represent the combination of two spectrally distinct transient channels whose spectral profiles correspond to those of the more transient multimodal peaks (red curves). While this is entirely plausible, our data lack sufficient temporal resolution to determine the validity of this hypothesis.

**Mechanistic interpretations**—High temporal frequency-biased responses (Derrington & Fuchs, 1978; Enroth-Cugell & Robson, 1966; Dawis, Shapley, Kaplan, & Tranchina, 1984; Hochstein & Shapley, 1976) and adaptation effects have been observed in magnocellular layers of primate retina (Baccus & Meister, 2002; Chander & Chichilnisky, 2001) and LGN (Solomon, Peirce, Dhruv, & Lennie, 2004). That a similar high temporal frequency-biased asymmetry to those observed in our masking and adaptation experiments occurs relatively early in visual processing suggests that the temporal whitening effects observed in Experiments 2–4 may be mediated by precortical mechanisms. Indeed, Dan, Atick, and Reid (1996) found direct evidence for temporal whitening of the kind observed in Experiment 2 in the responses of cat LGN neurons. When visually stimulated with naturalistic image sequences ( $\beta \approx -1$ ), temporal exponents derived from the firing rates of these cells were found to be approximately flat ( $\beta \approx 0$ ) between 3 and 15 Hz. Critically, these response-derived exponents became positively skewed ( $\beta \approx 1$ ) when stimulated with temporally unbiased broadband noise ( $\beta \approx 0$ ). Given that the high temporal frequency-biased responses of LGN appear to be consistent with temporal equalization of natural image sequences suggests that temporal whitening of the kind observed in Experiments 2–4 may be mediated as early in the visual pathway as LGN. It is conceivable, moreover, that these thalamic temporal whitening effects may themselves be inherited from the high temporal frequency response bias observed in retinal ganglion cells (peaking at around 10–20 Hz; Lee et al., 1990).

While these various neurophysiological studies suggest an early, precortical site for the temporal whitening effects observed in Experiments 2–4, one should consider the possibility that our psychophysical results may represent the combined effects of multiple discrete stages of processing, each with their own optimal tuning characteristics. Cross-orientation masking with analogous temporal tuning properties to those observed in Experiment 4 have been observed in cat V1, an effect which is variously attributed to intra-cortical inhibition (Morrone et al., 1987), feedback (Allison et al., 2001), precortical saturation (Li, Peterson, Thompson, Duong, & Freeman, 2005; Li, Thompson, Duong, Peterson, & Freeman, 2006; Priebe & Ferster, 2006), and thalamo-cortical suppression (Freeman, Durand, Kiper, & Carandini, 2002; Priebe & Ferster, 2006). Furthermore, the dissociations we observe in both the spectral locations and shape of masking and adaptation functions suggests that simultaneous (Experiments 1 and 2) and successive (Experiment 3) manifestations of temporal whitening may reflect distinct mechanisms.

## Conclusions

We show that the human visual perceptual response to spatially or temporally broadband noise is strongly biased in favor of high-frequency components. Critically, our results indicate that these high-frequency perceptual biases are of a magnitude that appears to equalize the physical low spatial and temporal frequency biases observed photometrically in natural scenes (Experiments 1 and 2). These perceptual equalization phenomena, respectively referred to as spatial and temporal whitening, support the idea that the human visual system is optimized to minimize redundant low-frequency-biased information present in natural scenes, an effect that may serve to enhance metabolic efficiency (Barlow, 2001; Dan et al., 1996). The idea that the transient response bias in Experiment 2 reflects a whitening process optimized to the statistical properties of natural scenes is supported by our finding in Experiment 3, whereby this response bias decreases with increasing stimulus spatial frequency ( $>2$  cycles/deg), a result which is predicted by the greater proportion of high temporal frequency energy at higher spatial frequencies resulting from the small fixational eye movements evident during normal viewing (Rucci et al., 2007). These results indicate that the visual system whitens its spatial and temporal input to an extent that compensates for the  $1/f$  bias in natural scenes. However, natural scenes are  $1/f$  both spatially and temporally. The broadband whitening effects we report here are made in response to stimuli that are broadband in only spatial or temporal frequency and thereby avoid any interactions that might occur between spatial and temporal frequency channels in response to spatio-temporally broadband stimuli. We are currently investigating such interactions.

In our fourth experiment, we confirm our earlier finding of a perceptual asymmetry in temporal masking. Specifically, while sensitivity to 15-Hz targets is compromised by mask flicker rates  $>6$ –40 Hz, 1-Hz targets are robustly masked by low and high flicker rates alike. We show that this temporal asymmetry in the effects of temporal masking produces a transient bias in the underlying temporal frequency channels' responses (the perceptual correlates of which can be observed directly in Demonstration Movie 2). Importantly, the magnitude of this transient response bias is inversely proportional to the low temporal frequency bias present in natural scenes. In light of the strong correspondence between the equalizing effects of temporal masking and the point of subjective temporal equalization measured in Experiment 2, we propose that overlay temporal masking and temporal whitening may share a common mechanism.

In order to identify the source of this transient-biased perceptual response, we conducted an adaptation experiment (Experiment 5). This showed that temporal adaptation produces a similar asymmetry to that observed in our simultaneous masking experiment. Whereas low temporal frequencies preferentially adapt low-frequency tests, adapting to high frequencies

(>6 Hz) reduces sensitivity to low and high frequencies alike. This asymmetry in temporal adaptation corroborates our interpretation of our masking data (Experiment 4) that temporal frequency channels are not independent, and moreover that they interact in such a way that low temporal frequency channels receive noisy adaptive input from higher frequency-selective mechanisms, but not vice versa. We propose that it is this noisy interaction that drives temporal whitening via simultaneous temporal masking. By fitting these adaptation data within the constraints of our asymmetric noise model and convolving the result with a  $1/f$  temporal amplitude spectrum, we show that adaptation induces a weak transient bias in the visual system's response that may at least partially compensate for the naturally low temporal frequency bias present in natural scenes. In addition to redundancy reduction, the transient perceptual bias we observe in response to broadband temporal luminance modulation may serve to optimize visual sensitivity to the naturally less prevalent transient information critical for the perception of motion and other dynamically driven segmentation processes.

## Acknowledgments

We wish to thank Keith Langley for sharing his insights into signal and noise.

This research was supported by an Australian Research Council Discovery Project (APD) grant DP0774697, awarded to John Cass.

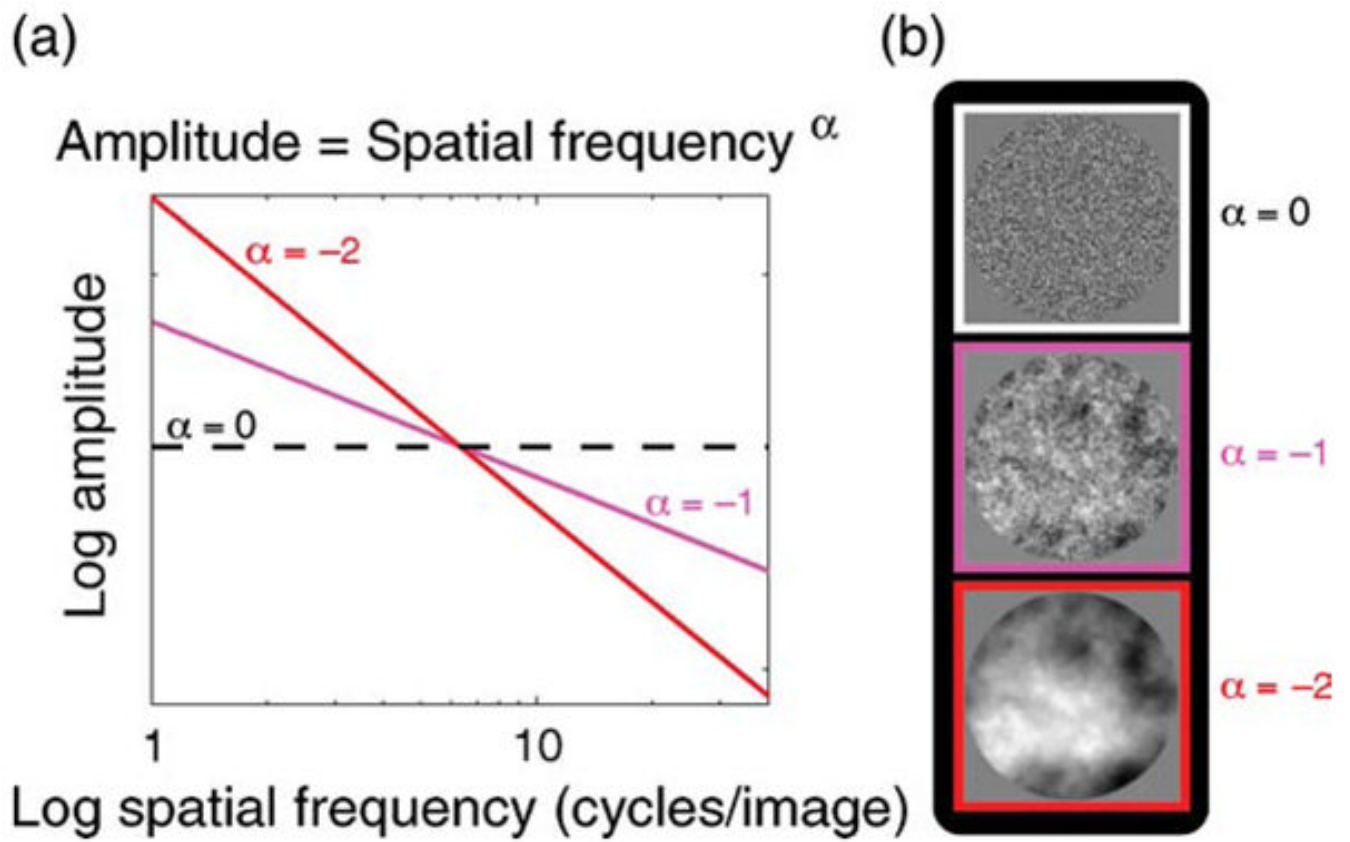
## References

- Alais D, Verstraten FA, Burr DC. The motion aftereffect of transparent motion: Two temporal channels account for perceived direction. *Vision Research* 2005;45:403–412. [PubMed: 15610746]
- Allison JD, Smith KR, Bonds AB. Temporal-frequency tuning of cross-orientation suppression in the cat striate cortex. *Visual Neuro-science* 2001;18:941–948.
- Anderson SJ, Burr DC. Spatial and temporal selectivity of the human motion detection system. *Vision Research* 1985;25:1147–1154. [PubMed: 4071994]
- Anderson SJ, Burr DC. Receptive field properties of human motion detector units inferred from spatial frequency masking. *Vision Research* 1989;29:1343–1358. [PubMed: 2635464]
- Atick JJ, Redlich AN. What does the retina know about natural scenes? *Neural Computation* 1992;4:196–210.
- Baccus SA, Meister M. Fast and slow contrast adaptation in retinal circuitry. *Neuron* 2002;36:909–919. [PubMed: 12467594]
- Barlow H. Redundancy reduction revisited. *Network* 2001;12:241–253. [PubMed: 11563528]
- Bex PJ, Dakin SC, Mareschal I. Critical band masking in optic flow. *Network* 2005;16:261–284. [PubMed: 16411499]
- Bex, PJ.; Langley, K. The perception of suprathreshold contrast and fast adaptive filtering; *Journal of Vision*. 2007. p. 1p. 1-23.<http://journalofvision.org/7/12/1/>
- Blakemore C, Campbell FW. Adaptation to spatial stimuli. *The Journal of Physiology* 1969a;200:11P–13P.
- Blakemore C, Campbell FW. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology* 1969b;203:237–260. [PubMed: 5821879]
- Boynton GM, Foley JM. Temporal sensitivity of human luminance pattern mechanisms determined by masking with temporally modulated stimuli. *Vision Research* 1999;39:1641–1656. [PubMed: 10343857]
- Brady N, Field DJ. What's constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns. *Vision Research* 1995;35:739–756. [PubMed: 7740766]
- Brady N, Field DJ. Local contrast in natural images: Normalisation and coding efficiency. *Perception* 2000;29:1041–1055. [PubMed: 11144818]
- Brainard DH. The Psychophysics Toolbox. *Spatial Vision* 1997;10:433–436. [PubMed: 9176952]

- Bredfeldt CE, Ringach DL. Dynamics of spatial frequency tuning in macaque V1. *Journal of Neuroscience* 2002;22:1976–1984. [PubMed: 11880528]
- Burton GJ, Moorhead IR. Color and spatial structure in natural scenes. *Applied Optics* 1987;25:157–170. [PubMed: 20454092]
- Campbell FW, Robson JG. Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology* 1968;197:551–566. [PubMed: 5666169]
- Cass J, Alais D. Evidence for two interacting temporal channels in human visual processing. *Vision Research* 2006;46:2859–2868. [PubMed: 16684555]
- Chance FS, Nelson SB, Abbott LF. Synaptic depression and the temporal response characteristics of V1 cells. *Journal of Neuroscience* 1998;18:4785–4799. [PubMed: 9614252]
- Chander D, Chichilnisky EJ. Adaptation to temporal contrast in primate and salamander retina. *Journal of Neuroscience* 2001;21:9904–9916. [PubMed: 11739598]
- Clifford CW, Webster MA, Stanley GB, Stocker AA, Kohn A, Sharpee TO, et al. Visual adaptation: Neural, psychological and computational aspects. *Vision Research* 2007;47:3125–3131. [PubMed: 17936871]
- Dan Y, Atick JJ, Reid RC. Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *Journal of Neuroscience* 1996;16:3351–3362. [PubMed: 8627371]
- Dawis S, Shapley R, Kaplan E, Tranchina D. The receptive field organization of X-cells in the cat: Spatiotemporal coupling and asymmetry. *Vision Research* 1984;24:549–564. [PubMed: 6740975]
- Derrington AM, Fuchs AF. Spatial and temporal frequency tuning in cat lateral geniculate nucleus [proceedings]. *The Journal of Physiology* 1978;282:45P–46P.
- Derrington AM, Fuchs AF. Spatial and temporal properties of X and Y cells in the cat lateral geniculate nucleus. *The Journal of Physiology* 1979;293:347–364. [PubMed: 501605]
- Dong DW, Atick J. Statistics of time-varying natural images. *Network* 1995;6:345–358.
- Durand S, Freeman TC, Carandini M. Temporal properties of surround suppression in cat primary visual cortex. *Visual Neuroscience* 2007;24:679–690. [PubMed: 17686200]
- Enroth-Cugell C, Robson JG. The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of Physiology* 1966;187:517–552. [PubMed: 16783910]
- Field DJ. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and Image Science* 1987;4:2379–2394.
- Field DJ, Brady N. Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research* 1997;37:3367–3383. [PubMed: 9425550]
- Foley JM, Boynton GM. Forward pattern masking and adaptation: Effects of duration, interstimulus interval, contrast, and spatial and temporal frequency. *Vision Research* 1993;33:959–980. [PubMed: 8506639]
- Freeman TC, Durand S, Kiper DC, Carandini M. Suppression without inhibition in visual cortex. *Neuron* 2002;35:759–771. [PubMed: 12194874]
- Georgeson MA. Over the limit: Encoding contrast above threshold in human vision. In: Kulikowski, JJ.; Walsh, V.; Murray, IJ., editors. *Vision and visual dysfunction: Vol. 5. Limits of visual perception*. London: Macmillan; 1990. p. 109–119.
- Georgeson MA, Harris MG. The temporal range of motion sensing and motion perception. *Vision Research* 1990;30:615–619. [PubMed: 2339514]
- Graham DJ, Chandler DM, Field DJ. Can the theory of “whitening” explain the center-surround properties of retinal ganglion cell receptive fields? *Vision Research* 2006;46:2901–2913. [PubMed: 16782164]
- Hess RF, Snowden RJ. Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research* 1992;32:47–59. [PubMed: 1502811]
- Hess RF, Waugh SJ, Nordby K. Rod temporal channels. *Vision Research* 1996;36:613–619. [PubMed: 8855005]
- Hochstein S, Shapley RM. Quantitative analysis of retinal ganglion cell classifications. *The Journal of Physiology* 1976;262:237–264. [PubMed: 994039]

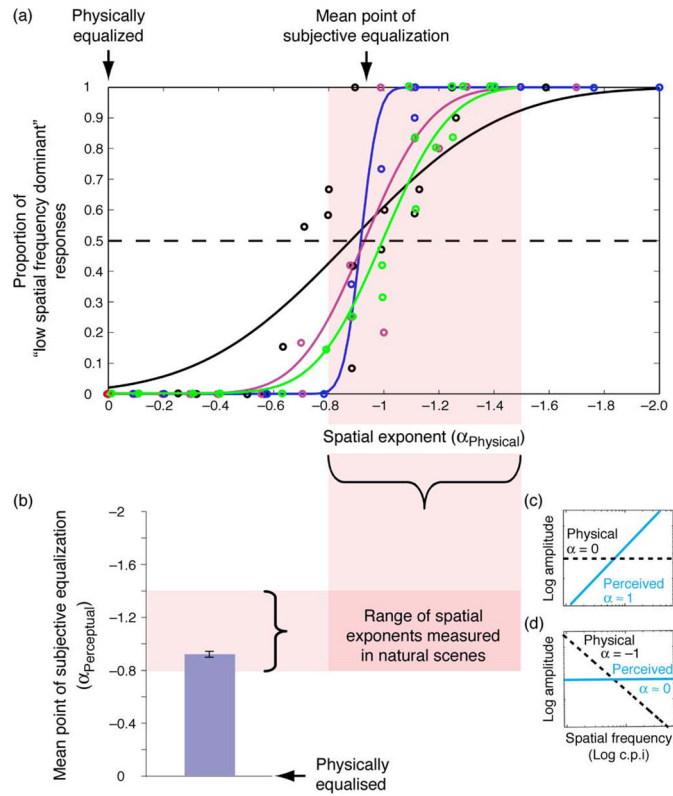


- Kohn A. Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology* 2007;97:3155–3164. [PubMed: 17344377]
- Kohn A, Movshon JA. Neuronal adaptation to visual motion in area MT of the macaque. *Neuron* 2003;39:681–691. [PubMed: 12925281]
- Langley K, Bex PJ. Contrast adaptation implies two spatiotemporal channels but three adapting processes. *Journal of Experimental Psychology: Human Perception and Performance* 2007;33:1283–1296. [PubMed: 18085943]
- Lee BB, Pokorny J, Smith VC, Martin PR, Valberg A. Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A, Optics and Image Science* 1990;7:2223–2236.
- Lehky SR. Temporal properties of visual channels measured by masking. *Journal of the Optical Society of America A, Optics and Image Science* 1985;2:1260–1272.
- Li B, Peterson MR, Thompson JK, Duong T, Freeman RD. Cross-orientation suppression: Monoptic and dichoptic mechanisms are different. *Journal of Neurophysiology* 2005;94:1645–1650. [PubMed: 15843483]
- Li B, Thompson JK, Duong T, Peterson MR, Freeman RD. Origins of cross-orientation suppression in the visual cortex. *Journal of Neurophysiology* 2006;96:1755–1764. [PubMed: 16855109]
- Meese TS, Holmes DJ. Spatial and temporal dependencies of cross-orientation suppression in human vision. *Proceedings of the Royal Society of London B: Biological Sciences* 2007;274:127–136.
- Meier, L.; Carandini, M. Masking by fast gratings; *Journal of Vision*. 2002. p. 2p. 293-301.<http://journalofvision.org/2/4/2/>
- Morrone MC, Burr DC, Speed HD. Cross-orientation inhibition in cat is GABA mediated. *Experimental Brain Research* 1987;67:635–644.
- Pantle A. Flicker adaptation—I. Effect on visual sensitivity to temporal fluctuations of light intensity. *Vision Research* 1971;11:943–952. [PubMed: 5133266]
- Priebe NJ, Ferster D. Mechanisms underlying cross-orientation suppression in cat visual cortex. *Nature Neuroscience* 2006;9:552–561.
- Rucci M, Iovin R, Poletti M, Santini F. Miniature eye movements enhance fine spatial detail. *Nature* 2007;447:851–854. [PubMed: 17568745]
- Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annual Review of Neuroscience* 2001;24:1193–1216.
- Snowden RJ, Hess RF. Temporal frequency filters in the human peripheral visual field. *Vision Research* 1992;32:61–72. [PubMed: 1502812]
- Solomon SG, Peirce JW, Dhruv NT, Lennie P. Profound contrast adaptation early in the visual pathway. *Neuron* 2004;42:155–162. [PubMed: 15066272]
- Srinivasan MV, Laughlin SB, Dubs A. Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London Series B: Biological Sciences* 1982;216:427–459.
- Tolhurst DJ, Tadmor Y, Chao T. Amplitude spectra of natural images. *Ophthalmic & Physiological Optics* 1992;12:229–232. [PubMed: 1408179]
- van der Smagt MJ, Verstraten FA, van de Grind WA. A new transparent motion aftereffect. *Nature Neuroscience* 1999;2:595–596.
- van Hateren, JH.; van der Schaaf, A. Temporal properties of natural scenes. *Proceedings of the IS&T/SPIE: Vol. 265. Human Vision and Electronic Imaging*; 1996. p. 139-143.
- Wang XJ, Liu Y, Sanchez-Vives MV, McCormick DA. Adaptation and temporal decorrelation by single neurons in the primary visual cortex. *Journal of Neurophysiology* 2003;89:3279–3293. [PubMed: 12649312]

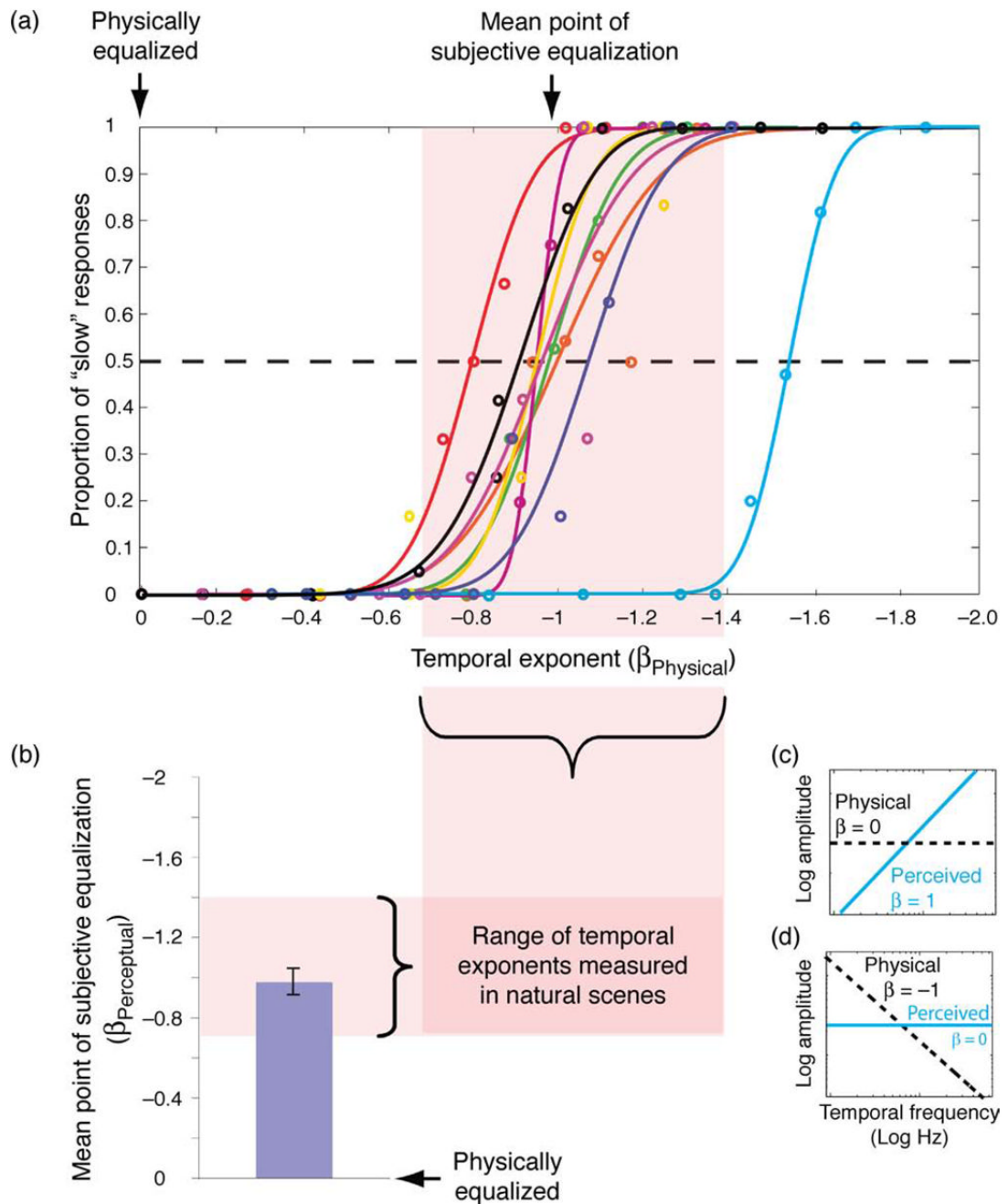


**Figure 1.**

(a) Spectral profiles of three different spatial exponents. (b) Examples of noise filtered with each of the spectral profiles depicted in panel a. White noise ( $\alpha = 0$ ) images are physically decorrelated, i.e., they possess an equal proportion of energy at all spatial frequencies. Natural images possess exponents ranging between  $-0.8$  and  $-1.5$ .

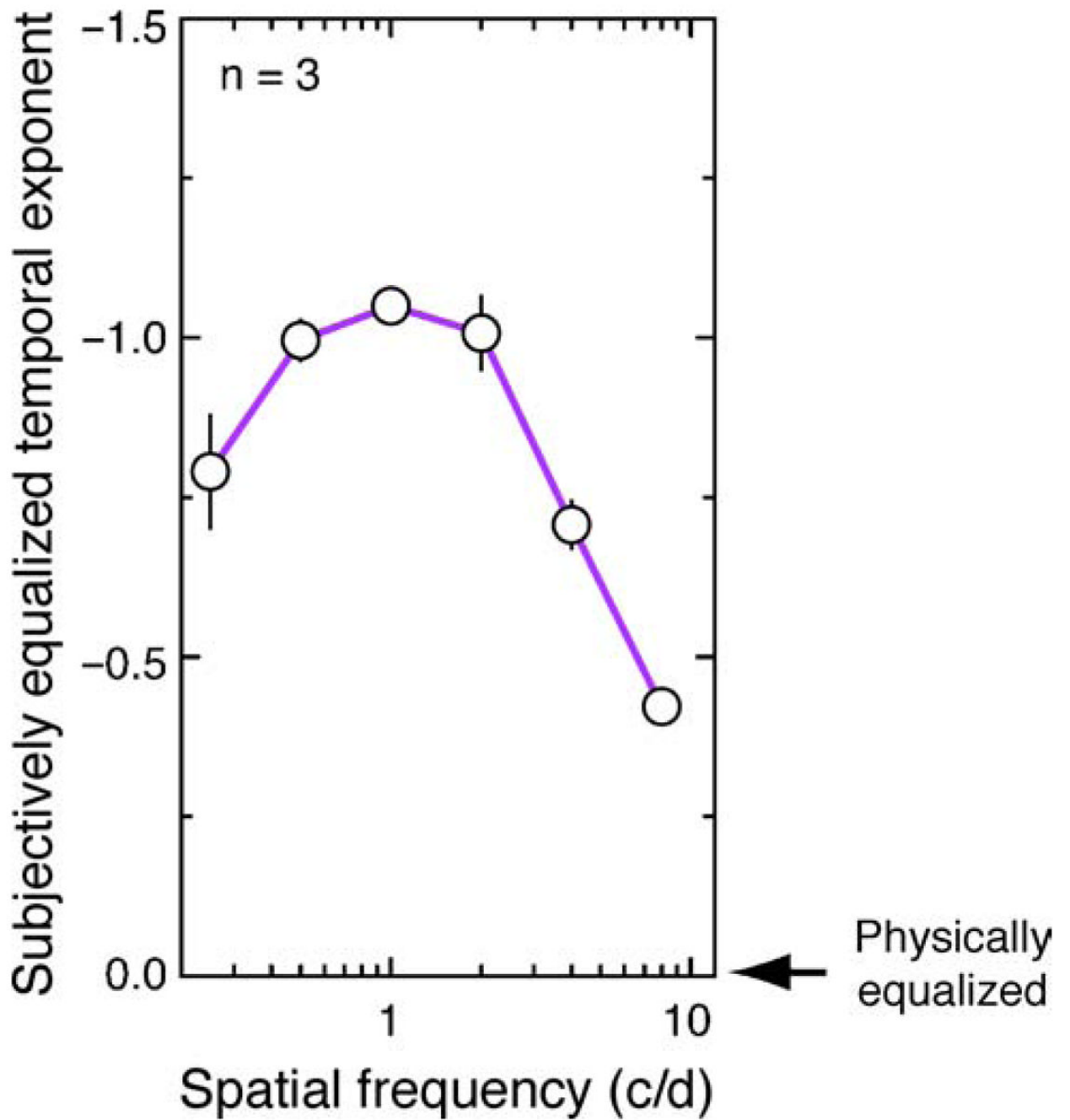


**Figure 2.** Perceptual estimates of spatial frequency content in exponentially filtered spatially broadband noise for four different observers. (a) Proportion of responses in which subjects reported stimuli to be low spatial frequency-biased measured as a function of the physical exponent of the stimulus. (b) Mean point of subjective equalization measured across subjects. Error bars represent inter-subject standard error. (c–d) Schematic depiction of the general finding that high frequencies are perceptually over-represented. Physically equalized stimuli ( $\alpha = 0$ ) are perceived to be strongly high frequency biased ( $\alpha \approx 1$ ) (c). Physical exponents ( $\alpha = -1$ ) are perceived to be approximately equalized (d).

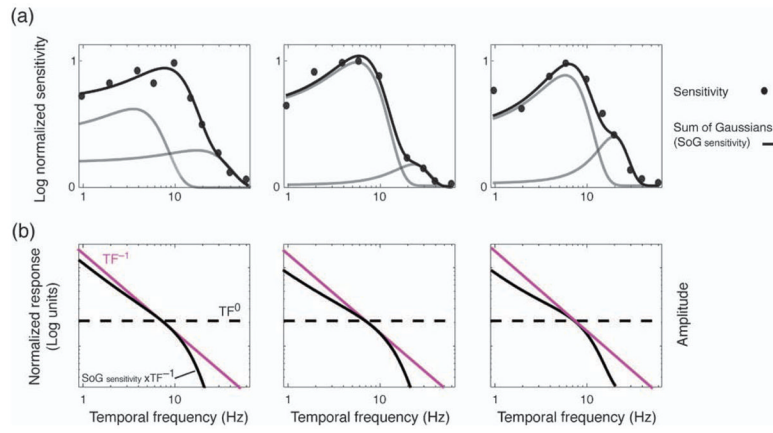


**Figure 3.**

Perceptual estimates of temporal frequency content in exponentially filtered temporally broadband (but spatially narrowband) noise. (a) Proportion of responses in which subjects reported stimuli to be low temporal frequency-biased measured as a function of the physical exponent of the stimulus. (b) Mean point of subjective equalization measured across subjects. Error bars represent inter-subject standard error. (c–d) Schematic depiction of general finding that high frequencies are perceptually over-represented. Physically equalized stimuli ( $\beta = 0$ ) are perceived to be strongly high frequency biased ( $\beta \approx 1$ ) (c). Physical exponents ( $\beta = -1$ ) are perceived to be approximately equalized (d).

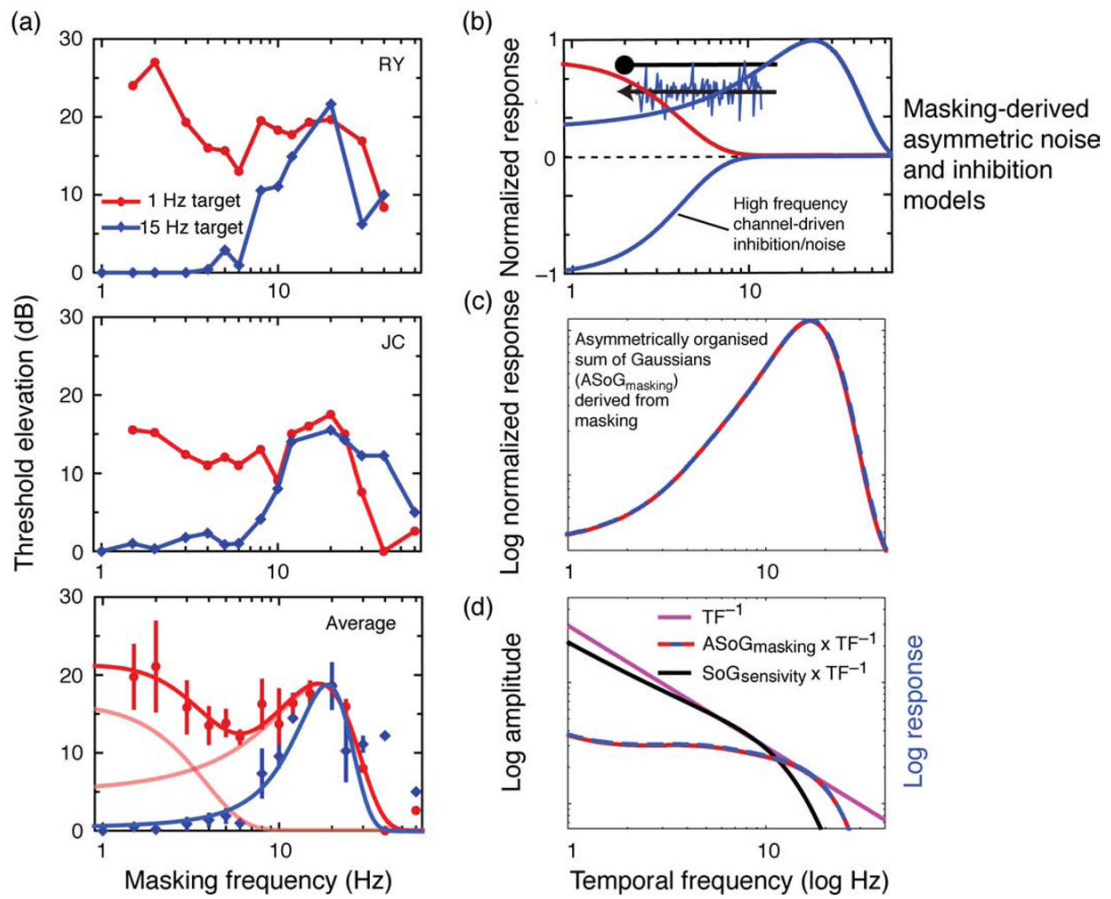


**Figure 4.** Effect of spatial frequency on the exponent at which broadband temporal frequency energy is perceived to be subjectively equalized. Error bars represent between subject mean standard error.



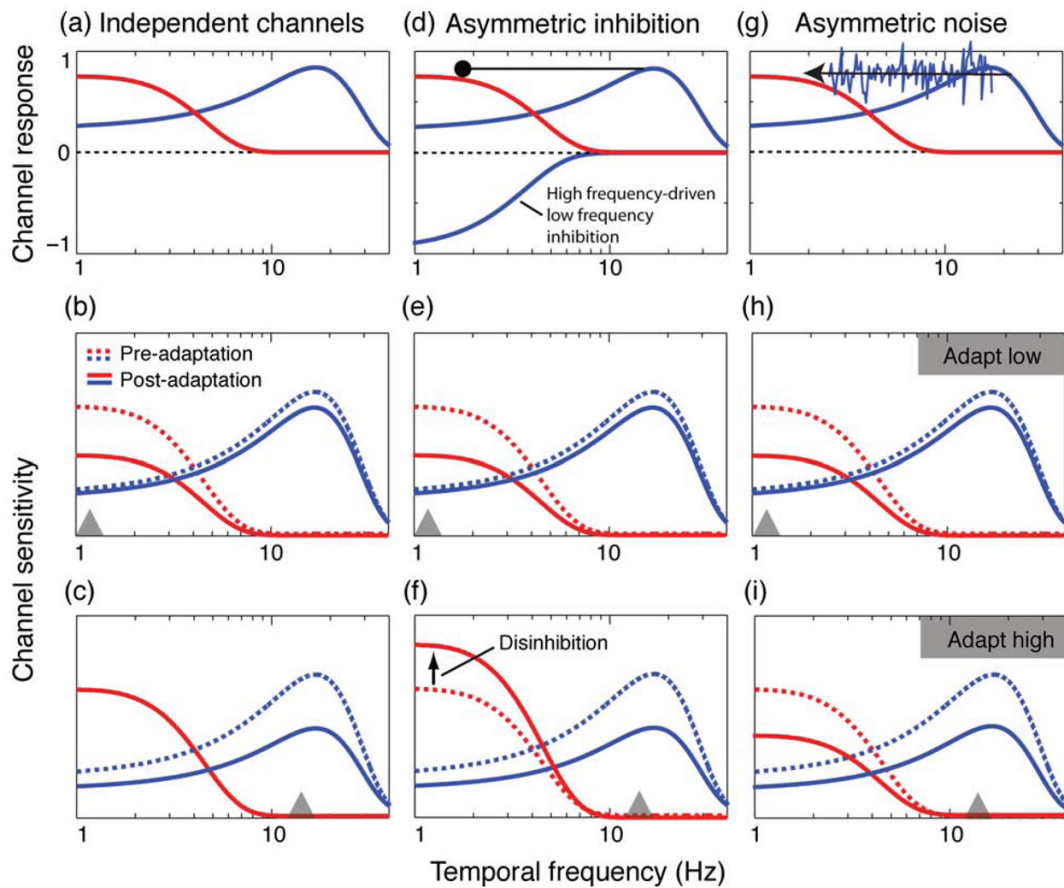
**Figure 5.**

(a) Log normalized modulation transfer functions for three separate observers. Black curve represents sum of two Gaussians (gray curves). (b) Inverse amplitude spectrum ( $\beta = 0$ ) (pink curve) convolved with sum of Gaussians (black curve) as depicted in panel a. Dashed line represents spectrally equalized temporal amplitude spectrum ( $\beta = 0$ ).



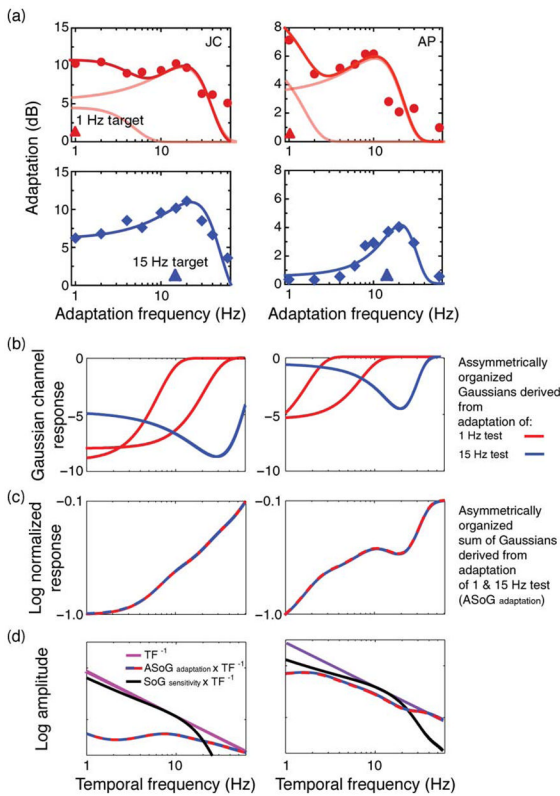
**Figure 6.**

(a) Results of temporal masking experiment using spatio-temporally narrowband isotropically filtered noise stimuli. Masking thresholds (y-axes) measured as a function of masking frequency (x-axis) derived using 1-Hz targets (red curves) and 15-Hz targets (blue curves). Bottom panel: depicts masking thresholds averaged across the two subjects; error bars represent inter-subject standard error; bold red and blue curves represent sum of two Gaussians (light red curves) and single Gaussian fitted to the averaged data. (b) Normalized response of masking-derived channels presented in the context of the asymmetric inhibition and asymmetric noise models (note that the negatively polarized function is a consequence of high-frequency channel-driven inhibition or noise). (c) Transient-biased sum of Gaussians derived from panel b. (d) Red-blue dashed curve: sum of Gaussians derived from panel c convolved with inverse amplitude spectrum ( $\beta = -1$ ) (pink curve). Black curve: modulation transfer function for JC (from Figure 5a).



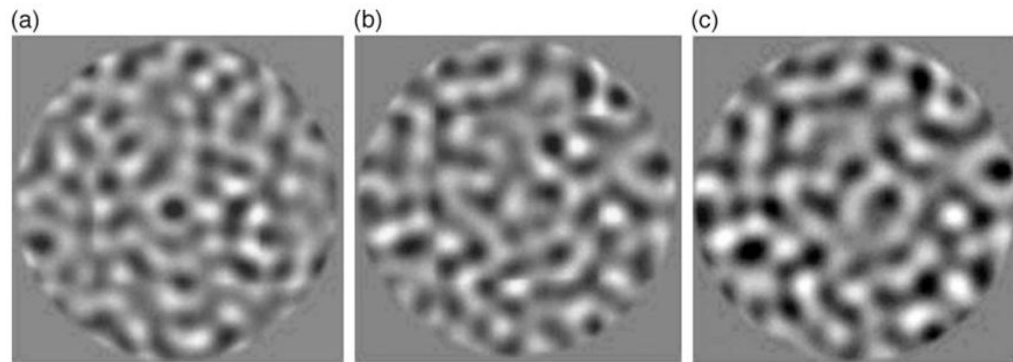
**Figure 7.** Predicted effects of adaptation for: independent channels (a–c); asymmetric inhibition (d–f); and asymmetric noise models (g–i). Dashed curves represent channel state prior to adaptation and solid curves represent predicted post-adaptive state. Gray triangles represent adaptation frequency. Each figure in the top row is a schematic depiction of each of the three models. Bottom two rows are schematic predictions.



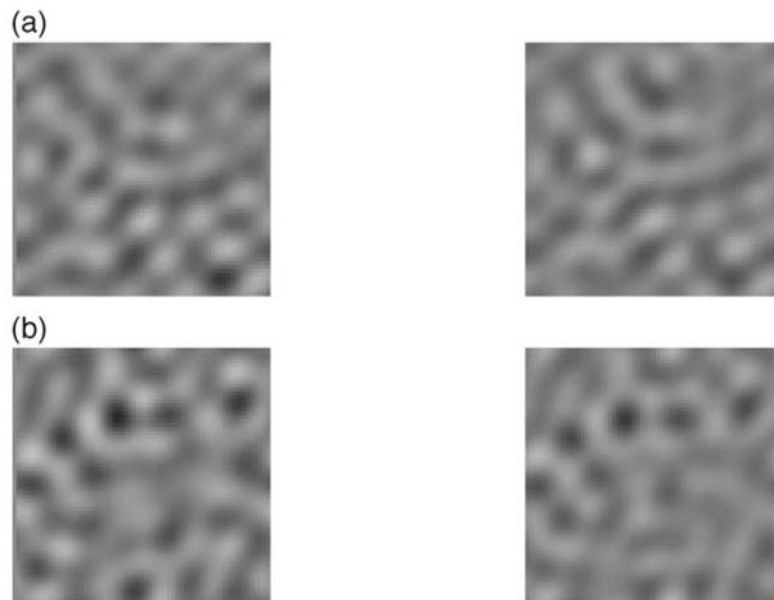


**Figure 8.**

(a) Results of temporal frequency adaptation experiment using spatio-temporally narrowband isotropically filtered noise stimuli. Each subject's data are represented in a separate column. Effect of different adaptation frequencies ( $x$ -axes) on detection thresholds ( $y$ -axes) to 1 Hz (red circles) and 15 Hz (blue diamonds) targets. Bold red curves represent fitted sum of three Gaussians (light red curves); bold blue curves represent a single fitted Gaussian. (b) Gaussians derived from panel a with spectral locations transposed according to asymmetric noise model (see text). (c) Sum of asymmetrically arranged Gaussians in panel b. (d) Red-blue dashed curve: Sum of asymmetrically arranged Gaussians from panel c convolved with inverse amplitude spectrum ( $\beta = -1$ ) (pink curve). Black curves represents modulation transfer functions depicted in panel a for each observer.

**Movie 1.**

Examples of temporally broadband filtered noise stimuli composed of different temporal exponents. Leftmost stimulus is composed of a physically equalized or “white” temporal amplitude spectrum ( $\beta = 0$ ). Temporal amplitude spectra of middle and rightmost stimuli are low frequency biased, with exponents of  $\beta = -1$  and  $\beta = -2$ , respectively. All stimuli are spatially narrowband.

**Movie 2.**

Demonstration of the temporal masking asymmetry. (a) The two stimuli on the left both contain  $\approx 10$ -Hz modulation at 50% contrast. The left most stimulus contains an additional  $\approx 1$ -Hz component (20% contrast). (b) The two stimuli on the right both contain  $\approx 1$ -Hz modulation, with the stimulus on the left containing an additional  $\approx 10$ -Hz component (20% contrast). While the 10-Hz component in panel b is clearly apparent, the 1-Hz component in panel a is far more difficult to detect.