The effect of notched noise on flicker detection and discrimination

Hannah E. Smithson Department of Psychology, Durham University, Durham, UK

G. Bruce Henning Institute of Ophthalmology, University College London, London, UK

Donald I. A. MacLeod Department of Psychology, University of California at San Diego, La Jolla, California, USA

Andrew Stockman Institute of Ophthalmology, University College London, London, UK

Flicker perception was investigated using two-alternative forced-choice detection and discrimination tasks with four different types of external noise: (1) broadband noise, (2) 5-Hz notched-noise—broadband noise with a 5-Hz band centered on the signal frequency removed, (3) 10-Hz notched-noise, and (4) no external noise. The signal was a burst of 10-Hz sinusoidal flicker presented in one of two observation intervals. In discrimination experiments, a pedestal—sinusoidal flicker with the same frequency, duration, and phase as the signal—was added to both observation intervals. With no noise, observers' performance first improved with increasing pedestal modulation, before deteriorating in accordance with Weber's Law, producing the typical "dipper" shaped plot of signal versus pedestal modulation. Noise affects performance, but the dipper effect persisted in each type of noise. The results exclude three models: the ideal-observer in which the pedestal improves performance by specifying the signal exactly; off-frequency-looking models in which the dipper depends on detection by channels tuned to temporal frequencies different from that of the signal; and strict energy detectors. Our data are consistent with signal processing by a single mechanism with an expansive non-linearity for near-threshold signal modulations (with an exponent of six) and a compressive "Weberian" non-linearity for high modulations.

Keywords: pedestal effect, noise, notched noise, flicker detection, flicker discrimination, off-frequency looking, uncertainty

Citation: Smithson, H. E., Henning, G. B., MacLeod, D. I. A., & Stockman, A. (2009). The effect of notched noise on flicker detection and discrimination. *Journal of Vision, 0*(0):1, 1–18, http://journalofvision.org/0/0/1/, doi:10.1167/0.0.1.

Introduction

Under certain conditions, the ability of a human observers to discriminate correctly which of two observation intervals contains a signal (sometimes called the target) can be improved by adding copies of the signal (usually called pedestals) to both intervals. This effect, known as the pedestal or dipper effect, runs counter to models in which the difference signal required for discrimination increases monotonically with background level as predicted by Weber's Law, in which the signal increases in proportion to the background level, or by the DeVries-Rose square-root law, in which it increases in proportion to the square-root of the background level.

The dipper effect is typically obtained in experiments in which the spatial and temporal properties of the signal and pedestal are matched in frequency, phase and orientation. The threshold-versus-contrast (TvC) function (in which the contrast [or modulation] of the signal corresponding to some percentage of correct responses is plotted against the pedestal contrast) exhibits a characteristic "dipper" appearance—as pedestal contrast increases from zero, performance first improves, and then deteriorates at higher pedestal levels (see, for example, Figure 1, below).

The earliest reports of the pedestal effect were for the discrimination of a flashed, uniform target superimposed on one of two spatially-separated flashed pedestals of the same size and duration (e.g., Barlow, 1962a, 1962b; Cornsweet & Pinsker, 1965; Whittle & Swanston, 1974), or for the discrimination of a grating presented on one of two temporally-separated gratings of the same spatial frequency and orientation (Campbell & Kulikowski, 1966).

The pedestal effect has received considerable attention in sensory research where it has been used as a means of investigating the suprathreshold properties of visual mechanisms. It has been used extensively in the spatial domain to investigate the response characteristics of channels or mechanisms that are differentially sensitive to spatial frequency and orientation (e.g., Bird, Henning, & Wichmann, 2002; Foley & Legge, 1981; Henning & Wichmann, 2007; Legge & Foley, 1981; Nachmias &

doi: 10.1167/0.0.1 Received June 17, 2008; published 0 0, 2009 ISSN 1534-7362 © ARVO

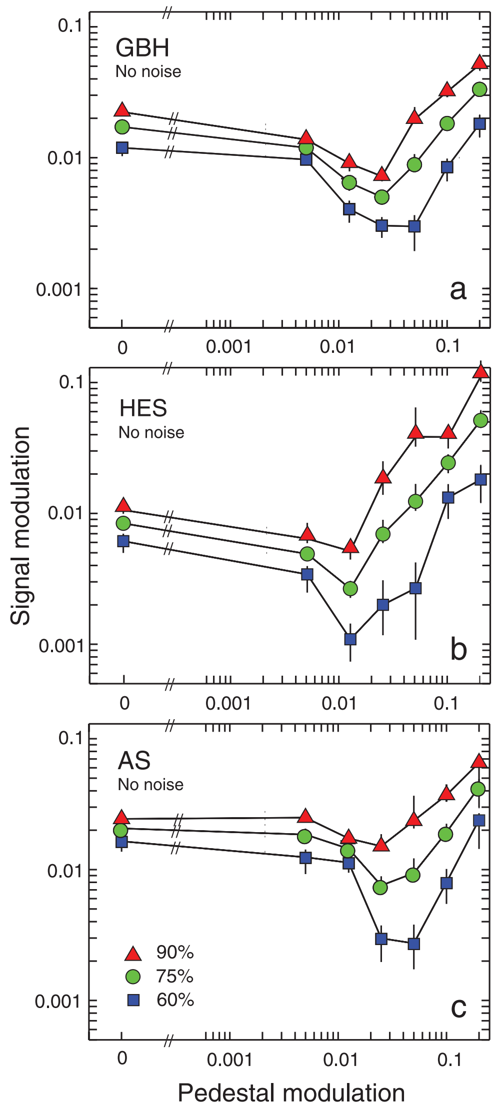

Figure 1. Data from the no-noise condition. Signal modulations corresponding to performance levels of 90% (red triangles), 75% (green circles), and 60% correct (blue squares) plotted as a function of the pedestal modulation. For each observer, the contours of constant performance were derived from Gumbell fits to the underlying psychometric functions at each pedestal level, based on at least five points, each of 100 observations (Wichmann & Hill, 2001a, 2001b). The logarithmic thresholds and their error estimates were converted to linear scales. The leftmost points were obtained with no pedestal. Vertical lines indicate ±1 standard deviation derived from the maximum likelihood fits to the psychometric functions. Observers: GBH (a), HES (b) and AS (c).

Sansbury, 1974; Wichmann, 1999; Yang & Makous, 1995). But it has also been used to investigate the properties of, for example, color and luminance mechanisms (e.g., Chen, Foley, & Brainard, 2000; Cole, Stromeyer, & Kronauer, 1990; Mullen & Losada, 1994; Switkes, Bradley, & De Valois, 1988); and ON- and OFF-channels (e.g., Bowen, 1995). The effect has also been reported with flickering or drifting gratings (e.g., Anderson & Vingrys, 2001; Boynton & Foley, 1999; Stromeyer, Kronauer, & Madsen, 1984) and with uniform flickering targets (Anderson & Vingrys, 2000; Stockman & MacLeod, 1985).

Here, we use the pedestal effect in the temporal domain to investigate the response characteristics of mechanisms sensitive to temporal frequency.

So, by what mechanisms might increasing the pedestal contrast in both intervals improve performance? A number of more or less plausible explanations of the pedestal effect have been proposed (see Solomon, 2009 for recent review). Broadly speaking, they can be put into four categories:

(1) The effect is the result of a specific nonlinear transducer function (e.g., Foley & Legge, 1981; Legge & Foley, 1981; Nachmias & Sansbury, 1974), such that the early part of the function is accelerating and the later part decelerating. The accelerating portion generates the dipper, because the difference in output between signal-plus-pedestal and the pedestal alone is larger than the difference in output between the signal alone and no signal, while the decelerating portion produces Weber's Law by compression. In some versions, the deceleration is produced by a divisive gain control (e.g., Boynton & Foley, 1999; Foley, 1994).

(2) The effect is due to a specific nonlinear transducer function combined with a signal-dependent internal noise (e.g., Green, 1967; Kontsevich, Chen, & Tyler, 2002), such that the accelerating nonlinearity produces the dipper at low pedestal levels, while the noise produces Weber's Law at high levels.

In both categories (1) and (2), the pedestal effect is assumed to be a characteristic of a single mechanism.

(3) Perhaps the most radical proposal is that the effect, in spatial vision at least, is due not to the characteristics of a single mechanism but to the pooled characteristics of many mechanisms with non-linear transducer functions that are insufficient in themselves to produce substantial dippers. The dipper is assumed to be produced by the recruitment of mechanisms that are mistuned away from the signal and pedestal as the pedestal contrast first increases (Goris, Wichmann, & Henning, 2009; Henning & Wichmann, 2007). We refer to these models as the "off-frequency-looking" model.

(4) Another, now somewhat discredited proposal (e.g., Bowen, 1995; Yang & Makous, 1995), is that the pedestal, because it is a copy of the test, reduces uncertainty about the frequency, phase, timing, and location of the signal thereby producing improved performance and the dipper (Pelli, 1985).

Here we use a similar strategy to Henning and Wichmann (2007) to evaluate these models, but applied in the temporal rather than spatial domain. We measured thresholds for detecting Hanning-windowed bursts of 10-Hz sinusoidal flicker in one of two temporal intervals containing pedestals of the same temporal frequency, phase and duration as a function of pedestal contrast (i.e., TvC functions). Measurements were made under four conditions of external noise: 1) broadband noise, 2) 5-Hz "notched" noise—the same broadband noise from which a 5-Hz band of noise centered arithmetically on the signal frequency had been removed, 3) 10-Hz notched noise and 4) no external noise. Comparisons among the no-noise and noise conditions, allow us to evaluate the different models proposed to account for the dipper effect. If off-frequency looking is important in producing the dipper, then the use of notched-noise should minimize the contributions of off-frequency channels and thus destroy the dipper (as Henning and Wichmann (2007) found in the spatial domain). If, on the other hand, uncertainty reduction is important, then, for an ideal observer (for whom the pedestal defines the signal frequency precisely), changing the notch width of the noise should not affect the observer's performance when the pedestal is present.

Because the dipper persists in notched temporal noise, and because performance depends on the notch width, our results are inconsistent with off-frequency looking in the temporal domain and with uncertainty reduction as characterized by the signal-known-exactly (SKE) ideal-observer. Instead, our data can be accounted for by assuming a single channel with an appropriate nonlinear transducer function. Following the early proposals of Delboeuf (1873) and Fechner (1860), we develop a simple nonlinear transducer function that describes our entire data set. The development of this model is described in the final section of the paper. This modeling suggests that the dipper effect cannot be characterized by a simple energy detector. Instead, the required transducer has a steeply-rising threshold non-linearity with an exponent of about six (i.e. three times that of a simple energy detector).

Methods

Subjects

Two males (aged 50 and 64) and one female (aged 30) participated in this study. The study conforms to the standards set by the Declaration of Helsinki, and the procedures were approved by local ethics committees at University College London.

Procedure

We used a two-interval forced-choice task. On each trial, noise and pedestals (if used) were both presented in two 1-second long observation intervals separated by a 500-millisecond pause. The signal was added to one of the observation intervals of each trial. The interval that contained the signal was randomly selected, so that the signal was equally likely to be in the first or second interval. Following the second observation interval, there was a 1.5 second response interval during which the observers indicated, by pressing keys, which interval they thought had contained the signal. Auditory signals indicated the beginning of each observation interval and the start of the response interval. Feedback was provided by a fourth auditory signal that indicated which observation interval had contained the signal. Psychometric functions of at least five points of 100 observations each were obtained in blocked sessions relating the percentage of correct responses to the amplitude of the signal for each pedestal level and notch width. These measurements were obtained in four conditions of external noise: 1) no noise; 2) broadband noise, 3) 5-Hz "notched" noise, 4) 10-Hz notched noise. The order of sessions was counterbalanced within observers.

Apparatus

Flickering stimuli were presented on an LED-based photo-stimulator that allows fine control of the luminance of bright uniform fields up to high temporal frequencies (Pokorny, Smithson, & Quinlan, 2004; Puts, Pokorny, Quinlan, & Glennie, 2005). The output of the LEDs was controlled via an M-Audio soundcard, housed in a G3 Macintosh computer. A circular test field, comprised of light from four LEDs (with peak outputs at 460, 516, 558, and 660 nm), had an annular surround, comprised of light from a second set of four LEDs with peak outputs at the same wavelengths. The test field subtended 2 degrees of visual angle and the annular surround subtended 8 degrees. To minimize the contrast at the border between the central and surround fields, each of the surround LEDs in turn was perceptually matched to the center LED having the same wavelength composition. The relative levels of the four central LEDs were chosen such that the fields were metamers of the equal-energy spectrum, and appeared approximately achromatic. In this study, the luminances of the surround LEDs were held constant, and the four LEDs illuminating the central test field were modulated in-phase to produce variations in luminance. The mean luminance of both the surround and the central field was 30 cd/m^2, which was sufficient to guarantee rod saturation.

Specification of stimuli

The LED spectra were measured with a telescopic spectroradiometer (Gamma Scientific, San Diego, CA) and used in conjunction with estimated cone sensitivities (Stockman & Sharpe, 2000) to calculate the ratio of the outputs of the component LEDs required to produce a light metameric to equal energy white. The relation between the intensities specified by the program and those produced by the diodes was established with a radiometer (UDT Instruments, Orlando, FL). A linearizing look-up table was then created to generate a mapping from the level requested in software to the luminance output of each LED. The system calibrated in this way should allow accurate luminance modulations with a resolution of 16.5 bits per channel up to about 100 Hz (Puts et al., 2005). The temporal waveforms were generated digitally and loaded to a buffer (wavetable) using the CoreAudio commands in Mac OS X.

Each stimulus had a duration of 1 second, which corresponded to 44100 samples at the sampling rate of the soundcard. All temporal waveforms were first generated in software using MATLAB. The 10-Hz signals and pedestals were generated as simple sinusoidal waveforms. The noise waveforms were defined as linear combinations of sinusoids from a set whose frequencies were equally spaced at 1 Hz intervals up to 100 Hz. At each frequency, the amplitudes of both sine- and cosine-phase sinusoids were randomly selected from a Gaussian distribution of zero mean and fixed variance. Broadband noise of this sort is sometimes called Fourier-series band-limited white Gaussian noise. The variance of the Gaussian distribution is proportional to the mean noise-power density of the noise and we describe below how we chose the appropriate variance. Notched noise was produced by removing either a 10- or 5-Hz band of components from a region centered arithmetically on the 10-Hz signal and pedestal frequency. The signals, in the frequency domain, were then transformed to the time domain.

We generated 100 noises in each noise class (meeting the criteria set out below). The noises were stored and, for each observation interval of our two-alternative forced-choice task, we randomly chose a noise from the appropriate class, each member of which was equally likely to be chosen. The waveform that was displayed was constructed by summing the appropriate signal, pedestal, and noise waveforms, multiplying the resultant waveform by a raised cosine (Hanning) window, rounding, and integerizing the windowed stimulus. The signal and pedestal were always in phase and in cosine phase with the peak of the window.

Calibration

To check the characteristics of the stimuli, a nominally sinusoidally flickering luminance was produced by the diodes and examined with the photometer. The photometer produced an electrical signal that followed the luminance input without loss up to about 100 Hz. We examined the harmonic content of a 10-Hz (nominally) sinusoidal flicker by sending the electrical output of the photometer through a wave analyzer (HP 35080A). This established that the stimulus was effectively sinusoidal since its second and third harmonic distortion products were negligible.

The photometer and the wave analyzer were also used to establish the characteristics of the flickering Gaussian noise. One-second long examples of the broadband noise and the 10-Hz notched noises were generated in MATLAB, rounded, integerized, displayed as repeating luminance waveforms through the diodes, and observed at the wave analyzer as the electrical signals from the photometer.

We chose the appropriate variance for the generation of the Gaussian noise by considering two related criteria: First we inspected the output from the diodes in response to broadband noise and increased the variance until the waveform was only very occasionally limited (clipped) by the maximum or minimum output; second, with the chosen value of the variance, we looked at the frequency spectrum of the notched noise using the wave analyzer. Notch depth is adversely affected either by excessive clipping (produced by too large a variance) or by insufficient dynamic range in the numerical representation prior to digital-to-analogue conversion (produced by too small a variance). For each noise sample we used, we confirmed that our 10-Hz notch had a stop-band in which the noise-power density was at least 35 dB below the noise-power density in the pass-band. A similar analysis of the 5-Hz notches was precluded by the finite bandwidth (1-Hz at half-power) of the narrowest filter in the wave analyzer. The mean root-mean-squared (r.m.s.) contrast of the 100 broadband noise samples used was 0.198, with a standard deviation of 0.008.

Results

Data obtained in the absence of external noise

The psychometric functions relating the percentage of correct responses to the logarithm of the depth of signal modulation were fit with Gumbel functions using the maximum-likelihood procedure of Wichmann and Hill (2001a, 2001b). Estimates of the modulation depths corresponding to 60%, 75%, and 90% correct responses, together with estimates of the variability associated with each estimate, were determined from these fits.

Figure 1 presents conventional threshold vs. pedestal functions, called threshold vs. contrast plots or TvC plots. Each panel shows, for a different observer, the signal modulation (or ripple ratio) corresponding to three different performance levels—90% correct (red triangles), 75% correct (green circles), and 60% correct (blue squares)—each as a function of the pedestal modulation; no external masking noise was used. Where larger than the data points, vertical lines indicate approximately ±1 standard deviation. The results are broadly similar for the three observers, and the pattern of results is roughly similar across the different performance levels: the contours of constant performance—which we refer to loosely as "thresholds"—first fall as the pedestal modulation increases from zero, and reach minima well below the "threshold" modulation depth obtained with no pedestal, before rising again. The minima are located at pedestal modulations just above the "threshold" modulations obtained with no pedestal. Comparable "dipper" shapes have been found in many other analogous experiments.

In spatial vision, the depth of the dipper and the location of its minimum depend on performance level: the dipper exhibited for low performance levels is deeper and occurs at higher pedestal levels than the dipper exhibited for higher performance levels (Bird et al., 2002; Goris, Wagemans, & Wichman, 2008; Wichmann, 1999). Similarly, for flicker, we find that the maximum improvement with added pedestal modulation is greater at lower than at higher performance levels and tends to occur at slightly higher pedestal modulations. The change in shape with performance level reflects the slopes of the underlying psychometric functions relating percentage correct to the logarithm of signal modulation, which are steepest at low pedestal levels, where the performance level contours are closely spaced, and most shallow in the vicinity of the dip, where the performance contours are most widely separated. The performance contours become more closely spaced once again on the rising portions of the curves where the pedestals mask the discrimination of the signal roughly in accordance with Weber's Law.

Data obtained with external noise added

We next consider the same detection and discrimination experiment performed in the presence of the three types of noise: 1) broadband, white Gaussian noise, 2) 5-Hz notched noise—the same broadband noise from which a 5-Hz band of noise arithmetically centered on the signal frequency had been removed, and 3) 10-Hz notched noise.

The three panels of Figure 2 show, separately for each observer, the 75% performance contours in the same format as Figure 1—the signal modulation producing 75% correct as a function of pedestal modulation. The black symbols are from the broadband-noise condition, the dark gray symbols from the 5-Hz-notch condition, the light gray symbols from the 10-Hz-notch condition, and the open symbols, from Figure 1, are from the no-noise condition. Error bars indicate approximately ±1 standard deviation. For all three observers the results vary systematically with the noise masking condition. Two changes are apparent with increasing notch width. First, the performance for the detection of the signal alone (i.e., the leftmost points in Figure 2) improves. Second, the region of masking by the suprathreshold pedestals begins at lower pedestal levels—and the pedestal value at which the best performance occurs decreases slightly—as notch width increases.

In Figure 3, we present the 60%, 75% and 90% performance contours for all conditions of the experiment

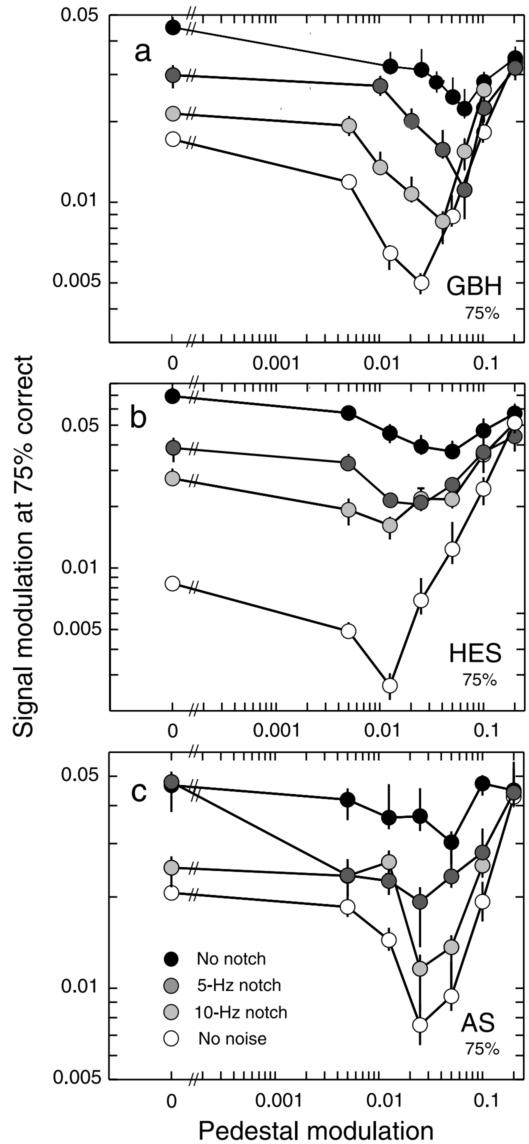

Figure 2. Signal modulations corresponding to 75% correct performance plotted as a function of pedestal modulation. Four different masking conditions are shown: data obtained with broadband white Gaussian noise (black circles), broadband noise from which a 5-Hz notch arithmetically centered on the signal frequency was removed (dark gray circles), broadband noise with a 10-Hz notch centered on the signal frequency (light gray circles), and with no noise (open circles, from Figure 1). Error bars were derived in a same way as for Figure 1. Observers: GBH (a), HES (b) and AS (c).

The three columns of plots show data for GBH (left), HES (middle) and AS (right). Plots in the top row show the results obtained with no external noise, and subsequent rows show data obtained with notched-noise maskers with a 10-Hz notch, notched-noise maskers with a 5-Hz notch and broadband noise maskers. Each plot compares data for the three performance levels: 60% contours (blue squares), 75% contours (green circles), and 90% contours (red

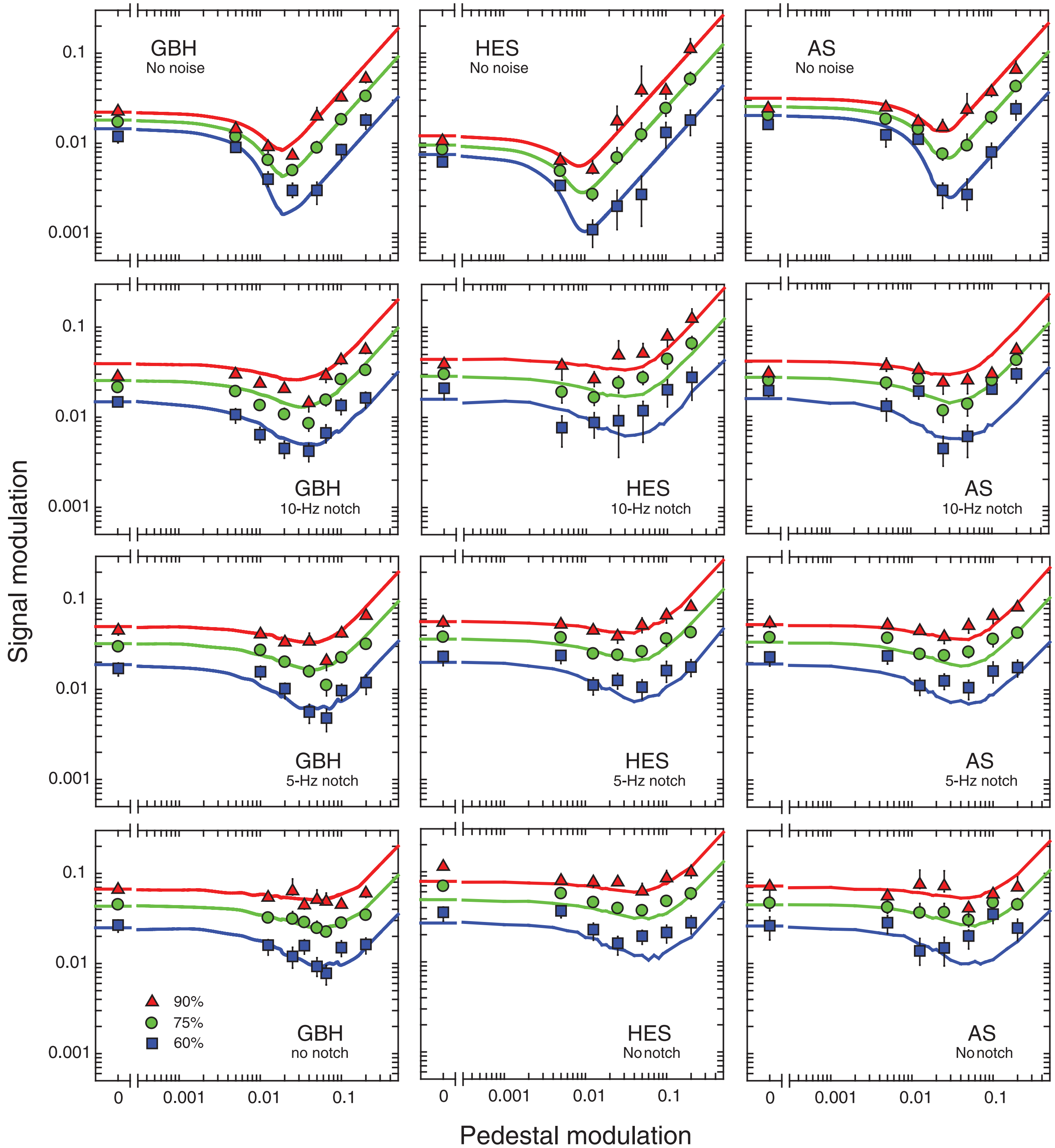

Figure 3. Data from four different masking conditions: no-noise (top row), broadband noise with a 10-Hz notch centered on the signal frequency (second row), broadband noise with a 5-Hz notch centered on the signal frequency (third row), and broadband white Gaussian noise (bottom row), for three observers: GBH, HES, AS. In each panel, signal modulations corresponding to performance levels of 90% (red triangles), 75% (green circles), and 60% correct (blue squares) are plotted as a function of the pedestal modulation. Smooth lines through the data are the best fitting curves from the non-linear transducer model of Equation 12. Details of simulation and fitting are provided in the text.

triangles). The solid lines through the data points show the best-fitting predictions of a model simulation described in the section "Development of a non-linear transducer model".

There is considerable variability among the observers, but many of the differences are due simply to differences in the observers' sensitivities. There are also several features of the data that are common to all three observers.

Associated with the improvement in performance level with increasing notch width for the detection of the signal alone (leftmost points), there is an accompanying increase in the separation between the performance contours defined by the data. Thus, increasing notch width causes the underlying psychometric functions to become steeper. One failure of our model predictions (shown by the continuous lines in Figure 3, and described later) is that the predicted contours at the detection threshold are approximately equally separated across the three noise conditions. As we discuss below, the changes we observe are also inconsistent with the uncertainty model. As in Figure 2 for the 75% contour, the contours at 60% and 90% also show that the extent of masking decreases with notch width, but that the facilitation—the dipper—persists across all conditions.

The characteristics of the contours at different performance levels apparent in the absence of external noise shown in Figure 1 are preserved in the presence of external noise: The size of the dipper depends on performance level, with the smallest improvement for the 90% performance contour and greatest improvement for the 60% contour. For GBH the dipper occurs close to, or slightly above, the detection threshold for the signal alone. This pattern is repeated for HES and AS, although the data are sometimes too noisy to locate the minima precisely. In general, in external noise conditions, the location of the dipper shifts to higher pedestal modulations compared to the location of the dipper in the absence of external noise.

"Threshold" signal modulation as a function of the combined strength of signal and pedestal

Some insight into the results can be obtained by plotting the signal modulation corresponding to some performance level against the combination of that signal modulation and the pedestal modulation (the modulations simply add in the combination because they are of the same frequency and phase). From the point of view of an observer, the task is either a detection or discrimination task, depending on the strength of the pedestal modulation. At low pedestal levels the task seems to the observer to be a detection task because the pedestal alone is never seen, whereas at high pedestal levels it seems to be a discrimination task—with the pedestal modulation alone in one interval and the signal-plus-pedestal modulation in the other.

Figure 4 shows the data for GBH from Figure 1 (no-noise condition) re-plotted with signal modulation as a function of signal-plus-pedestal modulation. In the top panel the signal modulation corresponds to 60% correct responses, in the center panel, to 75% correct, and in the

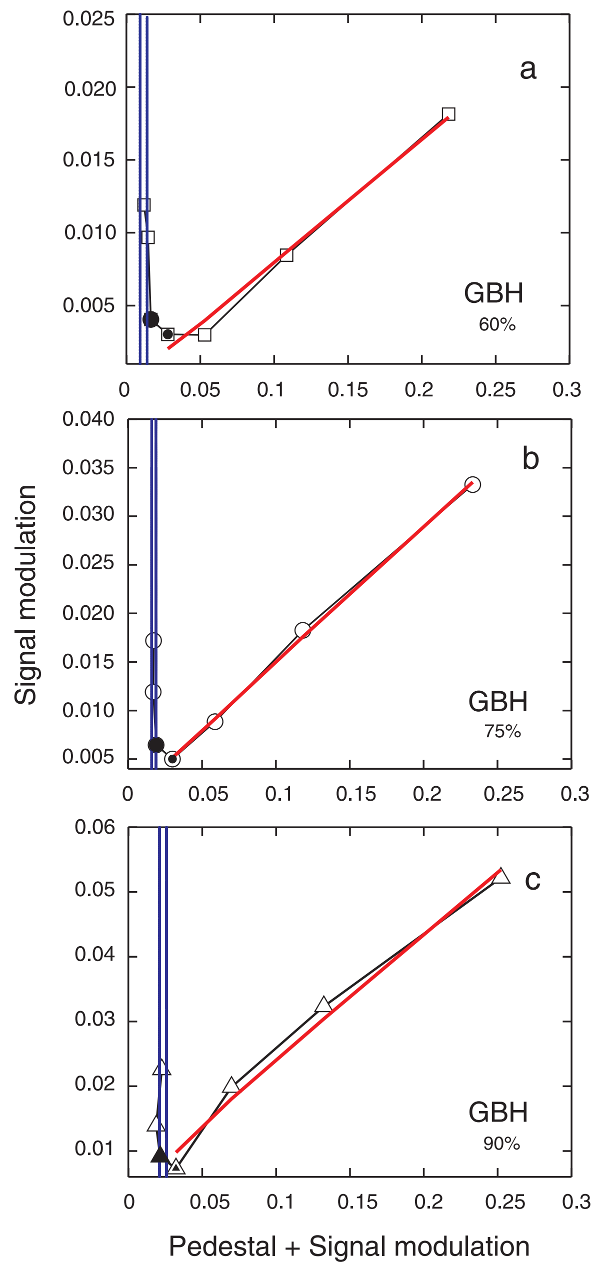

Figure 4. Data obtained in the no-noise condition (from Figure 1) for observer GBH plotted against different co-ordinates: each panel shows signal modulation as a function of signal-plus-pedestal modulation for 60% (a), 75% (b) and 90% (c) correct responses. The extended blue vertical lines in each panel mark the 95% confidence interval about the signal modulation required to achieve the appropriate performance level with zero pedestal modulation. The filled symbol in each panel marks the data point where the pedestal modulation alone is close to the 60% "threshold" and the partially filled symbol marks the data point where the pedestal alone is close to the 90% threshold. The red diagonal lines show the best (least squares) linear fit to the rightmost four points in each panel.

bottom panel, to 90% correct. The extended vertical lines (in blue) toward the left in each panel mark the 95% confidence interval around the signal modulation required to achieve that performance level in detecting the signal alone (i.e., with zero pedestal modulation). The approximate confidence intervals were obtained from the bootstrap procedure of Wichmann and Hill (2001a, 2001b). The red diagonal lines are the best (least squares) fits to the rightmost four points in each panel.

These graphs have several notable features. First, at low pedestal levels, the signal is discriminated when the signal-plus-pedestal modulation reaches the level at which modulation can be detected in the absence of the pedestal. For these pedestals of low modulation depth, the task is essentially a detection task; the pedestal alone is very rarely seen and the only interval with recognizable sinusoidal modulation comprises the pedestal modulation added to the signal modulation. The pedestal effect is produced because, in this region (i.e., for pedestal levels approaching the bottom of the dipper in Figure 1), it is the sum of the pedestal and signal modulation that produces the "threshold" stimulus; the signal modulation needed to reach the "threshold" decreases as the pedestal modulation increases and thus appears as the pedestal or dipper effect. This seems to be the case for all three performance thresholds.

Second, a narrow transition region begins at the point at which the modulation of the pedestal alone begins to be "seen". This transition region is delimited in each panel by the large and small filled symbols, which mark the approximate points on each curve at which the modulation of the pedestal reaches levels at which the pedestal *alone* should be detected with performance levels of 60% and 90%, respectively. In the transition region, the performance results from a mixture of detection-like trials, in which flicker with the temporal and spatial characteristics of the signal is seen in only one observation interval, and discrimination-like trials in which that flicker is seen in both intervals and the interval containing the more pronounced flicker (or flicker more like that of the signal) is chosen as having contained the signal. This region in Figure 4 is very small and corresponds, in effect, to the width of the psychometric function relating the percentage of correct responses to the depth of signal modulation in the absence of a pedestal.

Lastly, at higher pedestal levels, the signal modulation corresponding to a given performance level is proportional to the sum of signal and pedestal modulations. The red diagonal lines fitted to the upper three or four discrimination thresholds show the best (least squares) linear fit to the data in that region. The fitted function is of the form:

$$\Delta M = m(\Delta M + M) + c, \tag{1}$$

where m is the slope and c the intercept. All three observers produce results of the form of Figure 4 in the condition with no external noise. In all cases, the intercepts, c, are close to zero. The largest 95% confidence interval for the intercept, −0.030 to 0.029, was for the 90% performance contour for observer HES; all the remaining confidence intervals were within 0.01 of zero. This result is important, because it implies that in the regions in which performance can be described by Equation 1 it is governed, as in many discrimination tasks, by something like Weber's law; and it also means that the Weber fraction, $\Delta M/M$, can be extracted from the slopes of the linear fits. Rearranging Equation 1 with $c = 0$ gives:

$$\Delta M/M = m/(1 - m). \tag{2}$$

This is, of course, not a general finding, since not all contrast discrimination conforms to Weber's Law.

In Table 1, we summarize the fits of Equation 1 by tabulating the Weber fractions calculated using Equation 2 and the intercepts. The Weber fractions for modulation discrimination when the signal and pedestal have the same frequency and are in-phase correspond to the ratio of the signal modulation (at some "threshold" performance level) to the pedestal modulation. The average Weber fractions across the three observers are: 0.102, 0.177, and 0.247 for the 60%, 75%, and 90% performance contours, respectively.

For the conditions with noise, plots of the form of Figure 4 show similar characteristics to those obtained without noise. For example, data obtained for HES in the broadband noise condition are shown in Figure 5. In general, the interpretation of these plots for the conditions with external noise is slightly more difficult for two reasons: First the external noise introduces more variability (evident in the increased widths of the vertical blue lines giving the 95% confidence intervals for the

Observer	%	Weber Fraction	%	"Intercept"
GBH	60	0.091	60	−.00155
	75	0.139	75	.00101
	90	0.176	90	.00696
HES	60	0.089	60	.00007
	75	0.208	75	−.00077
	90	0.348	90	−.00089
AS	60	0.127	60	−.00414
	75	0.183	75	−.00161
	90	0.218	90	.00604
Average	60	0.102	60	−.00187
	75	0.177	75	−.00046
	90	0.247	90	.00403

Table 1. Weber fractions obtained at high pedestal levels corresponding to the percentage correct obtained in the no-noise condition and the corresponding intercepts of the least-squares linear fit to the rising sections of plots like those in Figure 4 for each observer and the average observer.

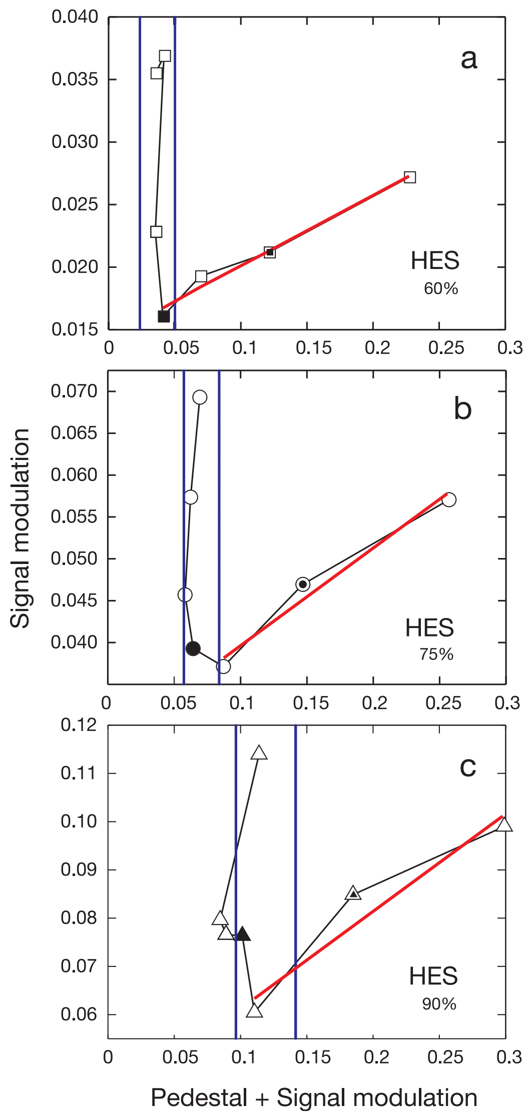

Figure 5. Data obtained in the broadband noise condition (from Figure 3) for observer HES plotted in the same format as Figure 4.

"thresholds" with zero pedestal levels), and second, the noise requires higher signal levels with the consequence that equipment constraints often preclude achieving high enough pedestal levels to ensure that only data on the rising part of the curve are included in the straight-line fit.

[Many disputes over the slope of the rising parts of TvC curves in spatial vision arise because of this problem (Wichmann, 1999)]. In the case of broadband noise, for example, only two of our observers (GBH and HES) appear to have three points on the rising part of their graphs. The best fitting lines again have intercepts close to zero and the average Weber fractions for these two observers are: 0.051, 0.096 and 0.149 for the 60%, 75% and 90% performance contours, respectively. That these values are smaller than those obtained in the no-noise condition could well be simply a result of our inability to generate pedestal levels that were high enough to get our observers into the linear range of the rising part of the graphs. One further effect in the broadband noise data for GBH and AS is a slight tendency for performance to improve beyond the transition region. This is also the case in spatial vision.

Of course, although instructive, the graphical representations in Figures 4 and 5 are essentially alternative representations of the TvC plots shown in the earlier figures. The vertical fall of ΔM with $\Delta M + M$ at low pedestal modulations is equivalent to a slope of -1 in the logarithmic TvC plots, whereas the linear growth of ΔM with $\Delta M + M$ with zero intercept at high pedestal modulations is equivalent to a slope of +1 in the logarithmic TvC plots. Neither graphical representation explains the data; any model that fits the underlying psychometric functions must have the characteristics of the data in both types of figure.

We now turn to explanatory models.

Discussion

The motivation behind these experiments was to further investigate the properties of the mechanisms that underlie flicker perception. Our approach has been to measure TvC functions under different conditions of external noise. These results enable us to do two things: first, to exclude some existing models of the pedestal effect based on off-frequency looking in the temporal domain and uncertainty reduction as characterized by the SKE ideal-observer; and, second, to develop a specific non-linear transducer model that can account for the entirety of our data. In this section, we discuss existing models.

Off-frequency looking models

The term off-frequency looking has been used to describe situations in which channels tuned to frequencies different from the signal frequency contribute to performance. In a recent study using spatially-varying stimuli and noise, Henning and Wichmann (2007) found that the dipper effect disappeared in notched broadband masking noise. They interpreted this as evidence that the pedestal effect results not from the characteristics of an individual spatio-temporal channel or mechanism, but rather from the way in which information is combined across diversely-tuned channels; i.e., observers rely on off-frequency looking in the region of the dipper (but see also Goris et al., 2009). However, contrary to these findings, we find that with temporally-varying stimuli the dipper effect survives in notched masking noise—a result that is inconsistent with models in which observers use information from channels tuned to different temporal frequencies.

Our results could be taken to imply that the activity of multiple mechanisms is not a necessary condition for the generation of the dipper, in which case they would pose a problem for off-frequency looking models, *in general*. However, off-frequency looking across spatial-frequency channels cannot be excluded by the results of our experiment.

The off-frequency looking model in spatial vision can be preserved by supposing that there is something fundamentally different between channels sensitive to temporal frequency and those sensitive to spatial frequency. One well-known difference is that there are fewer temporal frequency channels than spatial ones. Most estimates suggest two, or possibly three, flicker mechanisms (Boynton & Foley, 1999; Hess & Snowden, 1992; Levinson, 1960; Mandler & Makous, 1984; Roufs, 1974; Watson, 1986). By contrast, there are likely to be many spatial frequency channels (Blakemore & Campbell, 1969; Campbell & Robson, 1968; De Valois & De Valois, 1988; Graham & Nachmias, 1971; Henning, 1988; Henning, Hertz, & Hinton, 1981). Differences in channel numerosity alone, however, cannot explain why individual temporal frequency channels can sustain the full dipper effect, but individual spatial-frequency channels cannot. One possibility is that the temporal frequency channels have different underlying transducer functions, perhaps with a harder threshold nonlinearity, and perhaps mediated or limited by mechanisms earlier in the visual system than the emergence of spatial frequency channels.

It is also possible that the flicker response to our spatially-uniform flickering disc is mediated by a family of spatio-temporal channels optimally tuned to different (low) spatial frequencies. If the transducer functions of these spatial-frequency sensitive channels are similar to those tuned to the higher spatial frequencies investigated by Henning and Wichmann (2007), then the dipper that we find might also result from pooling across the spatial frequency domain.

Uncertainty reduction models

TvC functions measured under different conditions of external noise have also allowed us to evaluate explanations of the dipper effect based on uncertainty reduction. Such explanations suppose that the improvement in performance in the presence of the pedestal results from the pedestal improving the observer's knowledge of the characteristics of the signal (Burgess, 1985, 1990; Green & Swets, 1966; Pelli, 1985).

Uncertainty reduction models of the pedestal effect are typically assessed by comparing human performance with that of the ideal observer for a signal-known-exactly (SKE) (Burgess, 1985, 1990; Pelli, 1985). An ideal detection process takes advantage of full knowledge of the signal's waveform to filter out irrelevant frequencies and phases. Our unfiltered noise stimuli consist of sine and cosine components at 100 frequencies, each having identical independent Gaussian distributions of amplitude over trials so that the stimulus on a given trial defines a point in a 200 dimensional space. For a known signal, only one of the 200 dimensions is relevant, and the noise components for the other 199 dimensions can be ignored. But we show that this does not happen.

The ideal observer can be realized by using as the decision axis the output of a device that calculates the cross-correlation of the input (noise alone or signal plus noise) with a copy of the known signal (Green & Swets, 1966; van Trees, 1968). For sinusoidal signals, a cross-correlation mechanism is sensitive to only one component of the noise—that component having the same frequency and phase as the signal. Changing the width of a notch centered on the signal frequency has no effect on a cross-correlation receiver's performance and thus, if the cross-correlator is an adequate model of human behavior, changing notch widths should not affect human performance which, for both our notch widths, should be the same as having no external noise at all.

Our data, however, consistently show that performance varies systematically with the noise-masking condition. As described in the Results section, with increasing notch width the performance for the detection of the signal alone improves, and the underlying psychometric functions become steeper. These results are inconsistent with the behavior of the SKE ideal observer and indicate that the mechanism detecting the flicker responds to flickering noise of broad bandwidth rather than to a narrow band or to a single noise component like the ideal observer for the signal-known-exactly.

At the other extreme, if nothing is known about the frequency and phase of the signal, no such pruning of the stimulus space is possible; a stimulus located far from the origin in any direction is more likely to have originated from a signal plus noise rather than from noise alone, so the appropriate decision axis for an unknown signal is distance from the origin—the square root of the sum of squares of the sine and cosine amplitudes at all frequencies, which is monotonically related to the total flicker energy of the stimulus. We consider the energy detector subsequently.

Between the extremes of the SKE ideal observer and the energy detector there are many possible forms of uncertainty reduction—the coarse temporal-frequency discrimination of Mandler and Makous (1984), or the partition into 'agitation' as opposed to the luminance 'swell' visible at lower modulation frequencies (Roufs & Blommaert, 1981) suggest several—but their exploration is beyond the scope of this paper.

Non-linear transducer models

We argue that our results are broadly consistent with the behavior of a single mechanisms characterized either by a specific nonlinear transducer function (e.g., Foley & Legge, 1981; Legge & Foley, 1981; Nachmias & Sansbury, 1974), or by a specific nonlinear transducer function combined with a signal-dependent internal noise (e.g., Green, 1967; Kontsevich et al., 2002). In order to support this argument we next develop a specific non-linear transducer function that can account for the entirety of our data measured with and without masking noise.

Development of a non-linear transducer model

In this section, we use our entire data set, which is shown as symbols in Figure 3, to refine and develop a single-mechanism non-linear transducer model. However, rather than develop some arbitrary process, our strategy has been to constrain the modeling by starting with classical functions proposed by Delboeuf (1873) and Fechner (1860). The predictions of the developed model are shown by the continuous lines in Figure 3.

Data obtained without external noise

Fechnerian schemes for Weber's Law

We consider first the data obtained without external noise, the salient features of which are the approximate adherence to Weber's Law for large pedestal modulations, and the deviation from Weber's law characterized by the 'dipper' for near-threshold pedestal modulations.

Weber's Law characterizes the relation between stimulus magnitude, M (in our case, the pedestal modulation) and the difference in magnitude, ΔM (in our case, the added signal modulation) that is needed to make the combined modulation $M + \Delta M$ just noticeably different from M. In its simplest form, the Weber relation is $\Delta M = wM$, where the proportionality constant, w, is called the Weber fraction.

Fechner (1860) showed how the above form of Weber's Law could result from a logarithmic nonlinearity in the sensory response: On the assumption that all just-noticeable differences correspond to a constant difference in a sensory response ΔR, where $\Delta R = \Delta M / M$, Equation 3 follows by integration:

$$R(M) = \log_e(M) + C. \tag{3}$$

Equivalently, with $\log_e(M') = -C$ (where M' is the value of M for which $R(M) = 0$), Equation 3 becomes:

$$R(M) = \log_e(M/M'). \tag{4}$$

The black curve of Figure 6 shows this relationship, for $M' = 0.1$, with M on a linear scale in the upper panel and on a logarithmic scale in the lower panel. Now ΔR, the difference in R corresponding to a just noticeable stimulus difference $\Delta M = wM$, is always $\log_e(1 + w)$, independent of M, as shown in the derivation of Equation 5:

$$\begin{aligned} R(M + \Delta M) &= R[(1 + w)M], \\ &= \log_e[(1 + w)M/M'] \\ &= R(M) + \log_e(1 + w). \end{aligned} \tag{5}$$

Fechner did not provide a statistical account of Weber's Law applicable to forced-choice measures of discriminability. But if we make the standard assumption that, on

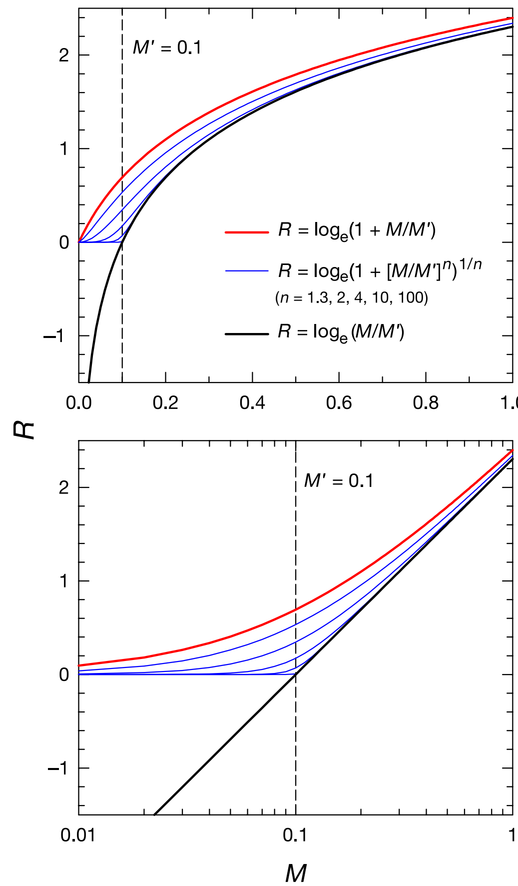

Figure 6. Response output (R) predicted as a function of modulation input (M) for various input-output schemes. Fechner's (1860) logarithmic input-output nonlinearity, which is given by $R = \log_e(M/M')$—see Equation 4—is shown for $M' = 0.1$ as the black continuous lines in each panel. Note that below $M = M'$, the response becomes negative, and is truncated in Equation 7. The modification of the nonlinearity by Delboeuf (1873) that keeps R positive, which is given by $R = \log_e(1 + M/M')$—see Equation 9—is shown for $M' = 0.1$ as the red continuous lines in each panel. Lastly, a series of functions of the form $R = \log_e(1 + [M/M']^n])^{1/n}$—see Equation 10—are shown as the blue continuous lines for $M' = 0.1$. From left to right $n = 1.3$, 2, 4, 10 and 100. As n increases, the transition at $M' = 0.1$ becomes increasingly abrupt (or hard). The lower panel is simply a semi-logarithmic version of the upper panel.

each observation interval of a trial, a logarithmically compressed neural signal deviates from its expected value $R(M) = \log_e(M/M')$ by the addition of Gaussian noise having a standard deviation, σ, independent of R (i.e., we assume the internal noise after the transducer is constant), then equal differences in R (and correspondingly equal fractional increases in M) will be detected with equal reliability whatever the starting value of R.

This provides a statistical and mechanistic "neo-Fechnerian" basis for Weber's Law. By this account, the Weber fraction w is set by the noise standard deviation σ, which has the same units as R and can be thought of as the equivalent root-mean-squared (r.m.s.) variation in the stimulus modulation from observation-interval to observation-interval, expressed as a fraction of the mean modulation M. The difference in R between two intervals with modulations M and $(1 + w)M$ is distributed with standard deviation $\sqrt{2}\sigma$ around its mean of $\log_e(1 + w)$, which is approximately w when w is small. Referring this to the cumulative Gaussian distribution, σ is equal to the Weber fraction w for a criterion of 76% correct 2AFC performance.

Generalizing Fechner: Hard threshold model

$R(M)$ as defined above decreases smoothly toward zero as the modulation M decreases to M'. But when M is less than M', R becomes negative, and it becomes increasingly negative without limit as M approaches zero (as indicated by the black curve of Figure 6). Fechner (1860) dealt with this unwelcome feature of the log transform by suggesting that the negative values of R correspond to 'unconscious sensations' that are all introspectively equivalent to one another, since none are consciously registered. As Fechner's contemporaries were quick to point out (e.g., Müller, 1878), a simple and natural alternative proposal is that the sensory response R simply remains zero for all $M < M'$. With this assumption, Fechner's log transform is truncated, replacing the negative values by zero (i.e., the lower-most blue line in Figure 6). The threshold modulation for eliciting a nonzero response, M', divides the response-modulation function into two regions. Below M' the response is zero, above M' it is positive and logarithmically compressed (though, approximately linear just above threshold where M is not much greater than M'):

$$R = \begin{cases} 0 & \text{for } M \le M' \\ \log_e(M/M') & \text{for } M > M', \end{cases} \tag{6}$$

or equivalently,

$$R = \max[0, \log_e(M/M')] \tag{7}$$

Just as the log transform provides a Fechnerian basis for Weber's Law, the threshold nonlinearity at M' in Equation 7 provides a Fechnerian basis for the dipper. All subthreshold modulations $M <= M'$ yield the same (zero) response, so pedestal and signal modulations that by themselves produce zero response can combine to produce a modulation that is discriminable from the (zero) response generated by the pedestal alone.

With the assumption introduced above, that the response R is contaminated by additive Gaussian internal noise of fixed variance, Equation 7 predicts performance in our experiments fairly well. Figure 7 shows the data for observer HES (replotted from the center panel of Figure 1) and the solid lines show the performance contours predicted by the model, and fitted with M' and σ as free parameters, estimated iteratively by using MATLAB's *fminsearch* function (based on the Nelder-Mead algorithm) to minimize the mean squared error of prediction in $\log_e(M)$. On each iteration, Equation 7 was used to evaluate the mean response, R_p for each experimental pedestal modulation, M_p (assuming the trial value for M'); the mean signal-plus-pedestal response required for criterion discrimination performance was then obtained as $R_{crit} = R_p + \sqrt{2}\sigma z_{crit}$, where z_{crit} is the standard normal deviate corresponding to the criterion percent correct, respectively 0.253, 0.674 and 1.282 for 60%, 75% and 90% correct responses. Equation 7 was inverted to determine the total modulation of signal and pedestal M_{crit} needed for the response R_{crit}, and then the required signal modulation M_s was obtained as $M_s = M_{crit} - M_p$.

Comparable fits were obtained for the other two observers, AS and GBH. The dippers predicted by this hard-threshold model tend to be a little deeper than ones observed, and the predicted psychometric functions with weak pedestals are slightly steeper than observed, as reflected in the tight spacing of the contours for different performance levels. But the transition from steep psychometric functions with weak pedestals to shallower ones with large pedestals is well

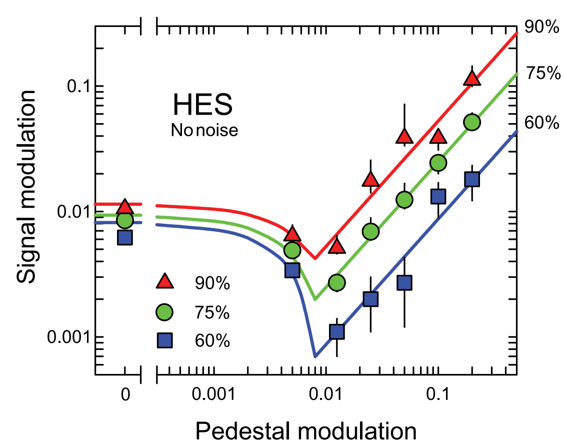

Figure 7. Data obtained in the no-noise condition (from Figure 1) for observer HES. Solid lines through the data are the best fitting curves from the hard-threshold model of Equation 7. Details of fitting are provided in the text.

predicted. The free parameters for the predictions of Figure 7 are $M' = 0.0075$, and $\sigma = 0.118$.

Generalizing Fechner: Small-signal linearity

In Fechner's time the dipper was neither experimentally recognized nor theoretically anticipated, but it was clear that Weber's Law had to be modified to accommodate small background stimulus magnitudes, since the simple formulation $\Delta M = wM$ implies discriminative capacity that improves without limit as background magnitude is decreased, contrary to observation. For many discrimination tasks, where no dipper is observed, a modified form of Weber's Law applies: the detectable stimulus increment has a progressive, linear relation to the combination of background stimulus magnitude M and a constant M' to which it is added:

$$\Delta M = w(M + M'). \tag{8}$$

In this formulation, M' is no longer the stimulus associated with zero response. In early discussions of intensity discrimination (Delboeuf, 1873), M' was regarded as the equivalent intensity of an effective background stimulus or 'intrinsic light,' always present and added to any external stimulus.

Delboeuf (1873) proposed an amendment to Fechner's logarithmic formula to make it consistent with this 'linear generalization' (Luce, 1959) of Weber's Law. This he did by simply substituting $(M + M')$ for M in Fechner's logarithmic formula, yielding:

$$\begin{aligned} R(M) &= \log_e[(M + M')/M'] \\ &= \log_e[1 + M/M']. \end{aligned} \tag{9}$$

The red curves in both panels of Figure 6 depict this relation. As can be seen, zero response to zero stimulus is still implied, but there is no sub-threshold dead zone.

Further generalization to incorporate intermediate (soft threshold) cases

The hard threshold of Equation 7 and the small-signal-linearity of Equation 9 can both be subsumed within a 'soft threshold' class of models that allow the gradient dR/dM to increase with various degrees of smoothness in the near-threshold range:

$$\begin{aligned} R &= \log_e\{[1 + (M/M')^n]^{1/n}\} \\ &= \log_e[1 + (M/M')^n]/n. \end{aligned} \tag{10}$$

Here the new parameter, n adjusts the "hardness" of the threshold while M' no longer necessarily corresponds to intrinsic light. Equations 7, 9 and 10 are asymptotically equivalent. The family of curves plotted with blue lines in Figure 6 show R as a function of M using Equation 10, for different values of the parameter n.

Relation to other non-linear transducer models

The three components of the models introduced here are also found in standard non-linear transducer models of the dipper effect (e.g., Foley & Legge, 1981; Nachmias & Sansbury, 1974; Wichmann, 1999): (i) a non-linear relation between stimulus modulation and some internal response, R; (ii) fixed internal noise added to R; and (iii) a decision mechanism. The shape of the predicted TvC function is strongly determined by the form of the response function provided the noise that limits the observers' behavior does not precede the nonlinearity (Lasley & Cohn, 1981; Peterson & Birdsall, 1953) and the dipper is typically modeled, as it is here, by assuming a response nonlinearity that is accelerative in the region of M'. In Equation 10, just as in Equation 7, M' is in that sense the "threshold" modulation, even though in Equation 10, a stimulus less than M' can elicit a response, and may be detectable without a pedestal if $w < 1$.

The response function and performance contours: How hard a threshold?

The two noted shortcomings of the predictions of Figure 7 can be alleviated by assuming a less than ideally-hard threshold through the appropriate choice of n in Equation 10. Softening the assumed threshold nonlinearity in Equation 10 rounds off and slightly elevates the bottom of the dipper, and also increases the predicted separation of the performance contours when the pedestal is sub-threshold or absent. With no pedestal, and small M, the contour spacing in a logarithmic plot is reduced when n is high, since the more accelerated the response function, the less is the change in stimulus modulation needed for a criterion change in response. But for pedestal modulations $M \gg M'$, where Weber's Law applies (at least asymptotically) for any n, the signal modulation must increase the natural log of the total modulation by $\sqrt{2}\sigma z_{crit}$, making the contour spacing wider and independent of n.

Equation 10 was used to fit the data for all subjects for the conditions where there was no external noise (assuming internal additive noise as before). All three parameters (M', σ and n) were varied iteratively for a best (least-squares) fit. The best fitting values of n were strikingly high (8, 7, and 5 for HES, AS and GBH, respectively), implying a very abrupt "threshold" nonlinearity. The large values of n that were required to fit the data illustrate the common failure of energy detectors ($n = 2$) to fit data of the sort we obtained (Wichmann, 1999). A value of $n = 2$ generates predictions that are obviously inaccurate (0.14 r.m.s. error in $\log_{10}$ modulation) in two respects: the dipper is clearly too shallow, and the spread between high and low criteria when no pedestal is present is too wide.

The exponents and fitting errors for the different subjects don't differ significantly, and on the basis of the pooled data the most likely exponent is about 6. But any value greater than about 3 gives a reasonably good fit to the data.

Data obtained with external noise added

A goal of these experiments was to analyze the detection process by investigating the ability of noise components of a range of different temporal frequencies to interfere with detection. To incorporate the effects of noise in our model, we assume that flicker is encoded as an energy-related quantity. Thus, information about the frequency and phase of the flicker is lost. Our limited ability in temporal-frequency discrimination (Mandler & Makous, 1984) supports this scenario—at least for relatively high frequencies, which appear as what Roufs and Blommaert (1981) call 'agitation' (as opposed to the luminance 'swell' visible at lower modulation frequencies).

Instead of deriving the flicker energy E from the amplitudes of the 200 Fourier components of the stimulus, it can be obtained directly as the sum of the squares of the time-varying excursion in relative luminance:

$$\Delta L_t / L_{ave} = (L_t - L_{ave}) / L_{ave}, \tag{11}$$

which, by Parseval's Theorem, is proportional to the sum of squares of the Fourier component amplitudes. Thus in the absence of external noise, E is proportional to the square of the signal amplitude, which is half the square of the modulation depth M in Equations 7, 9 and 10. Those equations can therefore be restated in terms of E/E' instead of M/M', with $E = M^2/2$ and the exponent n replaced by $n/2$, so that the best fitting exponent of $n = 6$ becomes $n = 3$ thus:

$$R = \log_e([1 + (E/E')^3]^{1/3}). \tag{12}$$

Whichever way the equation is expressed, the modulations are squared before the mean or sum is taken, and the sum is then subjected (approximately, in the near-threshold range, $E <= E'$) to a power-law (in this case a cubic) transform. But precise squaring of the deviations before integration is not critical to the predictions of energy-detection schemes, so long as the model prevents cancellation of positive and negative deviations. The energy detector is in this sense representative of a family of '*rectified transient*' detectors. When only external noise has to be considered, all detectors that base decisions on a monotonic function of energy perform equivalently (Lasley & Cohn, 1981; Peterson & Birdsall, 1953) and are effectively energy detectors. But if significant noise is added after the non-linearity the exponent in Equation 12 becomes critical. As noted above, linearity with energy (an exponent of 2 in Equation 10, or 1 in Equation 12 does not yield visually acceptable fits; linearity with modulation (halving the exponent) is even worse, predicting (in the absence of external noise) no dipper at all; but an energy detector with cubic response growth (Equation 12) gives a good account of our results without external noise. We consider next whether the energy-cubed model can predict performance with external noise as well.

Simulation methods

Thresholds in noise were estimated by simulating individual trials. The total noise energy E on any trial, expressed as a multiple of the expected energy of each noise component, is a sample from the chi-square distribution with the degrees of freedom equal to the number of independent noise components (e.g., 200 for the no-notch noise). When a signal or pedestal is present, the flicker energy is a sample from the non-central chi-square distribution, where the non-centrality parameter is the energy due to the sum of pedestal and signal. For each simulated presentation, the stimulus energy was generated by a random draw from the appropriate distribution, and the resulting response was obtained from Equation 12. Independent Gaussian internal noise of standard deviation σ was then added to the responses for each of the two presentations in a simulated 2AFC trial, and the decision was counted as correct if the response to the signal presentation was greater than to the no-signal presentation. We adopted the values for M' and σ that best fit the no-noise data for each subject, and a threshold hardness exponent $n = 3$ in accordance with Equation 12. Simulations were run on a range of test modulations spanning the full range of the psychometric function, with 10000 simulated trials per test modulation per pedestal, and the test modulations required for criterion performance were estimated by interpolation. For observers GBH, HES and AS, the best-fitting values of M' were 0.0165, 0.0081 and 0.0226 respectively, and the best-fitting values of σ were 0.1761, 0.2314 and 0.1961.

The critical band

Predictions for thresholds in noise depend on the bandwidth over which the noise energy is integrated. The simplest energy detector, where all noise frequencies are weighted equally, is implausible at the outset, since the highest frequencies in the 100-Hz noise band are invisible at our mean luminances, and although possibly present in neural responses (Hawken, Shapley, & Grosof, 1996; Lee, Sun, & Zucchini, 2007; Shady, MacLeod, & Fisher, 2004), are unlikely to contribute much masking. Moreover, the calculated performance assuming full sensitivity to all noise frequencies was vastly inferior to what we observed. This failure of prediction can be corrected by supposing that the energy computation is preceded by considerable filtering of the temporal luminance waveform, with a filter frequency response either completely determined by the temporal CSF or perhaps (if the signal is not completely unknown) also influenced by proximity to the signal frequency. This general structure (filter followed by rectification) is inherited from prior models, notably those of Goris et al. (2008), Rashbass (1970), and Roufs and Blommaert (1981).

For the model, we adopted a parametrically specified smooth passband shape, with a peak at the signal frequency (10 Hz) and adjustable width. A linear temporal filter was assumed, multiplying the effective modulation at frequency f by an attenuation factor:

$$A(f) = [(f/10)\exp(1-f/10)]^N. \tag{13}$$

The parameter N is the exponent of the rising low-frequency part of the modulation sensitivity function. The high frequency cutoff is also steeper for large N, so increasing N makes the passband narrower, preserving full transmission at 10 Hz. The bandwidth-narrowing exponent N was determined iteratively, with a complete simulation run for each iteration.

Values for N between 1 and 2 gave a good account of the data (Figure 3 shows the model predictions for $N = 1.4$, with a root mean square prediction error of 0.116).

The main features of the data are captured in the predictions shown in Figure 3, and the deviations from prediction are not very consistent across subjects. Appropriate choice of N yields good estimates of the overall amount of masking for the notch noises as well as for the broadband noise. The rightward shift of the dipper in the external noise conditions is also predicted (perhaps over-predicted) by the model, because threshold is set by total noise at the output, and the contribution of external noise to this total is greater for weak pedestals, where the gradient of the function relating energy to output (Equation 12) is steep. The required passband of the early filter is quite broad, ranging from about 3 to 25 Hz at half-height. This is quite comparable with the width of the temporal modulation sensitivity function, although the peak and width of that function vary considerably with the conditions of observation (Kelly, 1977; Robson, 1966). The filter bandwidth is, however, narrower than the bandwidth at the retinal output, which exceeds the psychophysical detection bandwidth (Lee et al., 2007). Evidently, most if not all of the visible noise is effective in reducing sensitivity to the test signal, as if the observer's decision is based on the total visibility-weighted flicker energy integrated over frequency.

Internal luminance noise

To provide an account of Weber's Law for flicker discrimination (pedestal-aided detection) we have assumed that internal noise is added to the neural representation of flicker after the nonlinear transform of Equation 10. This is equivalent to assuming in the Weber region that the internal noise before the transducer grows according to e^M. But internal noise may also be introduced in the form of random fluctuations in signals representing luminance—noise present in the input to the stages responsible for rectification and compressive nonlinearity. Although Figure 3 shows that such noise need not be invoked to provide an approximate account of the detection thresholds, it is expected a priori and indeed provides an important functional justification for threshold nonlinearity, as the nonlinearity would be helpful in rejecting small inputs that are likely to be due to internal noise at the input to the nonlinear stage (Morgan, Chubb, & Solomon, 2008; Simoncelli & Adelson, 1996).

The addition of small amounts of internal luminance noise does improve the hard threshold model, by appropriately increasing the range of uncertain vision (the separation of the performance contours) when the pedestal is absent or sub-threshold, thereby correcting one of the failings of that model seen in Figure 3. But too much internal luminance noise tends to obliterate the dipper, just as external noise does.

Summary

Psychometric functions relating the percentage of correct responses to the depth of modulation of a 10-Hz sinusoidally flickering stimulus were measured in standard two-alternative forced-choice experiments under various conditions of external noise. Our results are broadly inconsistent with uncertainty reduction and off-frequency looking explanations of the dipper effect and with a strict energy detector. Instead, they suggest that the dipper effect reflects some form of nonlinear transducer function within a single channel or mechanism. We have developed a specific non-linear transducer (starting with Fechner's early insight) that economically accounts for the entirety of our data set, with and without noise.

Acknowledgments

This research was supported by grants from the BBSRC and the Wellcome Trust awarded to AS, from the Leverhulme Trust and Wellcome Trust awarded to GBH, and from NIH (EY01711) to DIAM.

Commercial relationships: none.
Corresponding author: Hannah Smithson.
Email: hannah.smithson@durham.ac.uk.
Address: Department of Psychology, Durham University, Science Laboratories, South Road, Durham, DH1 3LE, United Kingdom.

References

Anderson, A. J., & Vingrys, A. J. (2000). Interactions between flicker thresholds and luminance pedestals. *Vision Research, 40,* 2579–2588. [PubMed]

Anderson, A. J., & Vingrys, A. J. (2001). Multiple processes mediate flicker sensitivity. *Vision Research, 41,* 2499–2455. [PubMed]

Barlow, H. B. (1962a). Measurements of the quantum efficiency of discrimination in human scotopic vision. *Journal of Physiology, 160,* 169–188. [PubMed] [Article]

Barlow, H. B. (1962b). A method of determining the overall quantum efficiency of visual discriminations. *Journal of Physiology, 160,* 155–168. [PubMed] [Article]

Bird, C. M., Henning, G. B., & Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A, 19,* 1267–1273. [PubMed]

Blakemore, C., & Campbell, F. W. (1969). Adaptation to spatial stimuli. *The Journal of Physiology, 200,* 11P–13P. [PubMed]

Bowen, R. W. (1995). Isolation and interaction of ON and OFF pathways in human vision: Pattern-polarity effects on contrast discrimination. *Vision Research, 35,* 2479–2490. [PubMed]

Boynton, G. M., & Foley, J. M. (1999). Temporal sensitivity of human luminance pattern mechanisms determined by masking with temporally modulated stimuli. *Vision Research, 39,* 1641–1656. [PubMed]

Burgess, A. E. (1985). Visual signal detection: III. On Bayesian use of prior knowledge and cross-correlation. *Journal of the Optical Society of America A, 2,* 1498–1507. [PubMed]

Burgess, A. E. (1990). High level visual decision efficiencies. In C. B. Blakemore (Ed.), *Visual coding and efficiency* (pp. 430–440). Cambridge: Cambridge University Press.

Campbell, F. W., & Kulikowski, J. J. (1966). Orientation selectivity of the human visual system. *Journal of Physiology, 187,* 437–445. [PubMed] [Article]

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551–556. [PubMed] [Article]

Chen, C.-C., Foley, J. M., & Brainard, D. H. (2000). Detection of chromoluminance patterns on chromoluminance pedestals: I. Threshold measurements. *Vision Research, 40,* 773–788. [PubMed]

Cole, G. R., Stromeyer, C. F., III, & Kronauer, R. E. (1990). Visual interactions with luminance and chromatic stimuli. *Journal of the Optical Society of America A, 7,* 128–140. [PubMed]

Cornsweet, T. N., & Pinsker, H. M. (1965). Luminance discrimination of brief flashes under various conditions of adaptation. *The Journal of Physiology, 176,* 294–310. [PubMed] [Article]

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision*. Oxford: Oxford University Press.

Delboeuf, J. R. L. (1873). Étude psychophysique. Recherches théoriques et expérimentales sur la mesure dessensations et spécialement des sensations de lumière et de fatigue *Mémoires couronnés et autres memoires publies par l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique* (vol. 23). Bruxelles: Hayez.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf and Hartel.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A, 11,* 1710–1719. [PubMed]

Foley, J. M., & Legge, G. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research, 21,* 1041–1053. [PubMed]

Goris, R. L., Wagemans, J., & Wichman, F. A. (2008). Modelling contrast discrimination suggests both the pedestal effect and stochastic resonance to be caused by the same mechanism. *Journal of Vision, 8*(15):17, 1–21, http://journalofvision.org/8/15/17/, doi:10.1167/8.15.17. [PubMed] [Article]

Goris, R. L. T., Wichmann, F. A., & Henning, G. B. (2009). A neurophysiologically plausible population-code for human spatial vision. *Journal of Vision,* in press.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11,* 251–259. [PubMed]

Green, D. M. (1967). Additivity of masking. *Journal of the Acoustical Society of America, 41,* 1517–1525. [PubMed]

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley & Sons.

Hawken, M. J., Shapley, R. M., & Grosof, D. H. (1996). Temporal-frequency selectivity in monkey visual cortex. *Visual Neuroscience, 13,* 477–492. [PubMed]

Henning, G. B. (1988). Spatial-frequency tuning as a function of temporal frequency and stimulus motion. *Journal of the Optical Society of America A, 5,* 1362–1373. [PubMed]

Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. Noise masks. *Journal of the Optical Society of America, 71,* 574–581. [PubMed]

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7*(1):3, 1–15, http://journalofvision.org/7/1/3/, doi:10.1167/7.1.3. [PubMed] [Article]

Hess, R. F., & Snowden, R. J. (1992). Temporal properties of human visual filters: Number, shape and spatial covariation. *Vision Research, 32,* 150–169. [PubMed]

Kelly, D. H. (1977). Visual contrast sensitivity. *Optica Acta, 24,* 107–129.

Kontsevich, L. L., Chen, C. C., & Tyler, C. W. (2002). Separating the effects of response non-linearity and internal noise psychophysically. *Vision Research, 42,* 1771–1784. [PubMed]

Lasley, D. J., & Cohn, T. E. (1981). Why luminance discrimination may be better than detection. *Vision Research, 21,* 273–278. [PubMed]

Lee, B. B., Sun, H., & Zucchini, W. (2007). The temporal properties of the response of macaque ganglion cells and central mechanisms of flicker detection. *Journal of Vision, 7*(14):1, 1–16, http://journalofvision.org/7/14/1/, doi:10.1167/7.14.1. [PubMed] [Article]

Legge, G., & Foley, J. M. (1981). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471. [PubMed]

Levinson, J. Z. (1960). Fusion of complex flicker II. *Science, 131,* 1438–1440. [PubMed]

Luce, D. R. (1959). On the possible psychophysical laws. *Psychological Review, 66,* 81–95. [PubMed]

Mandler, M. B., & Makous, W. (1984). A three channel model of temporal frequency perception. *Vision Research, 24,* 1881–1887. [PubMed]

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A 'dipper' function for texture discrimination based on orientation variance. *Journal of Vision, 8*(11):9, 1–8, http://journalofvision.org/8/11/9/, doi:10.1167/8.11.9. [PubMed] [Article]

Mullen, K. T., & Losada, M. A. (1994). Evidence for separate pathways for color and luminance detection mechanisms. *Journal of the Optical Society of America A, 11,* 3136–3151. [PubMed]

Müller, G. E. (1878). *Zur Grundlegung der Psychophysik. Kritische Beiträge [On the founding of psychophysics. Critical contributions].* Berlin: T. Grieben.

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039–1042. [PubMed]

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A, 2,* 1508–1532. [PubMed]

Peterson, W. W., & Birdsall, T. G. (1953). *The theory of signal detectability, Technical Report 13*. University of Michigan: Electronic Defense Group.

Pokorny, J., Smithson, H. E., & Quinlan, J. (2004). Photostimulator allowing independent control of rods and the three cone types. *Visual Neuroscience, 21,* 263–267. [PubMed]

Puts, M. J., Pokorny, J., Quinlan, J., & Glennie, M. (2005). Audiophile hardware in vision science; the soundcard as a digital to analog converter. *Journal of Neuroscience Methods, 142,* 77–81. [PubMed]

Rashbass, C. (1970). Visibility of transient changes in luminance. *The Journal of Physiology, 210,* 165–186. [PubMed] [Article]

Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America, 56,* 1141–1142.

Roufs, J. A. (1974). Dynamic properties of vision: IV. Thresholds of decremental flashes, incremental flashes and doublets in relation to flicker fusion. *Vision Research, 14,* 831–851. [PubMed]

Roufs, J. A., & Blommaert, F. J. (1981). Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. *Vision Research, 21,* 1203–1221. [PubMed]

Shady, S., MacLeod, D. I., & Fisher, H. S. (2004). Adaptation from invisible flicker. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 5170–5173. [PubMed] [Article]

Simoncelli, E. P., & Adelson, E. H. (1996). Noise removal via Bayesian wavelet coring. *Proceedings of 3rd IEEE Conference on Image Processing, 1,* 379–382.

Solomon, J. A. (2009). The history of dipper functions. *Attention, Perception, & Psychophysics, 71,* 435–443. [PubMed]

Stockman, A., & MacLeod, D. I. (1985). Visible beats from invisible flickering lights: Evidence that blue-sensitive cones respond to rapid flicker. *Perception, 14,* A18.

Stockman, A., & Sharpe, L. T. (2000). Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research, 40,* 1711–1737. [PubMed]

Stromeyer, C. F., III, Kronauer, R. E., Madsen, J. C., & Klein, S. A. (1984). Opponent-movement mechanisms in human vision. *Journal of the Optical Society America A, 1,* 876–884. [PubMed]

Switkes, E., Bradley, A., & De Valois, K. K. (1988). Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings. *Journal of the Optical Society of America A, 5,* 1149–1162. [PubMed]

van Trees, H. L. (1968). *Detection, estimation, and modulation theory*. New York: Wiley.

Watson, A. B. (1986). Temporal sensitivity. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (vol. 1, pp. 6-1–6-43). New York: Wiley.

Whittle, P., & Swanston, M. T. (1974). Luminance discrimination of separated flashes: The effect of background luminance and the shapes of t.v.i. curves. *Vision Research, 14,* 713–719. [PubMed]

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination*. Unpublished D. Phil., University of Oxford.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness-of-fit. *Perception and Psychophysics, 63,* 1293–1313. [PubMed] [Article]

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics, 63,* 1314–1329. [PubMed] [Article]

Yang, J., & Makous, W. (1995). Modelling pedestal experiments with amplitude instead of contrast. *Vision Research, 35,* 1979–1989. [PubMed]

AUTHOR QUERY

AUTHOR PLEASE ANSWER QUERY

AQ1: Please update the publication status of this reference.

END OF AUTHOR QUERY